

SOFTWARE: Patent.pm
; SEQ ID NO 407
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -18...-1
US-09-978-360A-407

Query Match 100.0%; Score 1626; DB 11; Length 302;
Best Local Similarity 100.0%; Pred. No. 2,4e-160;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIICSVFSAYIILCCWAGLPICLATCLDHHPTGSRPTVGPPLHFSYGS 60
DB 1 MKAPGRVLIIICSVFSAYIILCCWAGLPICLATCLDHHPTGSRPTVGPPLHFSYGS 60
QY 61 SVDPGKPLVREPCRSQAVSSSGOMLGSGIAEIDSAECVFRNNOAPTVGFEDVQGRST 120
DB 61 SVDPGKPLVREPCRSQAVSSSGOMLGSGIAEIDSAECVFRNNOAPTVGFEDVQGRST 120
QY 121 LRVVSHTSVPLLRNYSHYFQKARDTLNMGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
DB 121 LRVVSHTSVPLLRNYSHYFQKARDTLNMGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
QY 181 VYFTEEMMAYCQDIFODETGKRRSGSFLSTGFTMLALCEIIVVYGWVSDSYCR 240
DB 181 VYFTEEMMAYCQDIFODETGKRRSGSFLSTGFTMLALCEIIVVYGWVSDSYCR 240
QY 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAAPRSARHFTTEKAVFSRMAKKRPIVFAHPSWR 300
DB 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAAPRSARHFTTEKAVFSRMAKKRPIVFAHPSWR 300
QY 301 TE 302
DB 301 TE 302

RESULT 2

US-10-315-664-75
; Sequence 75; Application US/10315664
; Publication No. US20030203377A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bougueleret, L.
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal
; FILE REFERENCE: GENSET.050CP3
; CURRENT APPLICATION NUMBER: US/10/315.664
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US/09/599,360
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: 60/113,686
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/141,032
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/469,099
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ.ID NOS: 123
; SOFTWARE: Patent.pm
; SEQ ID NO 75
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -18...-1
US-10-315-664-75

Query Match 100.0%; Score 1626; DB 14; Length 302;
Best Local Similarity 100.0%; Pred. No. 2,4e-160;

Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIICSVFSAYIILCCWAGLPICLATCLDHHPTGSRPTVGPPLHFSYGS 60
DB 1 MKAPGRVLIIICSVFSAYIILCCWAGLPICLATCLDHHPTGSRPTVGPPLHFSYGS 60
QY 61 SVDPGKPLVREPCRSQAVSSSGOMLGSGIAEIDSAECVFRNNOAPTVGFEDVQGRST 120
DB 61 SVDPGKPLVREPCRSQAVSSSGOMLGSGIAEIDSAECVFRNNOAPTVGFEDVQGRST 120
QY 121 LRVVSHTSVPLLRNYSHYFQKARDTLNMGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
DB 121 LRVVSHTSVPLLRNYSHYFQKARDTLNMGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
QY 181 VYFTEEMMAYCQDIFODETGKRRSGSFLSTGFTMLALCEIIVVYGWVSDSYCR 240
DB 181 VYFTEEMMAYCQDIFODETGKRRSGSFLSTGFTMLALCEIIVVYGWVSDSYCR 240
QY 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAAPRSARHFTTEKAVFSRMAKKRPIVFAHPSWR 300
DB 241 EKSHPSVPHYFEKGRIDECOMYLAEHQAAPRSARHFTTEKAVFSRMAKKRPIVFAHPSWR 300
QY 301 TE 302
DB 301 TE 302

RESULT 3

US-09-895-298-66
; Sequence 66; Application US/09895298
; Publication No. US20030078405A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 47 Human Secreted Proteins
; FILE REFERENCE: P2035P1
; CURRENT APPLICATION NUMBER: US/09/895,298
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/591,16
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: PCT/US99/29950
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/113,006
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 60/112,809
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ.ID NOS: 231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (237)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-895-298-66

Query Match 99.8%; Score 1622; DB 10; Length 302;
Best Local Similarity 99.7%; Pred. No. 6,2e-160;
Matches 301; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIICSVFSAYIILCCWAGLPICLATCLDHHPTGSRPTVGPPLHFSYGS 60
DB 1 MKAPGRVLIIICSVFSAYIILCCWAGLPICLATCLDHHPTGSRPTVGPPLHFSYGS 60
QY 61 SVDPGKPLVREPCRSQAVSSSGOMLGSGIAEIDSAECVFRNNOAPTVGFEDVQGRST 120
DB 61 SVDPGKPLVREPCRSQAVSSSGOMLGSGIAEIDSAECVFRNNOAPTVGFEDVQGRST 120
QY 121 LRVVSHTSVPLLRNYSHYFQKARDTLNMGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
DB 121 LRVVSHTSVPLLRNYSHYFQKARDTLNMGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 180

Qy	181	VYTFPERMAYCDQI FQDETGNRRSGSFLSTGWFMTLIALTLCSEIVVYGWVSQSYCR	240
D6	181	VYTFPERMAYCDQI FQDETGNRRSGSFLSTGWFMTLIALTLCSEIVVYGWVSQDYCR	240
Qy	241	EKSHPSVPYPHYPEKGRLDECOMYLAEQA PRSAHRPTTEKAVFSRNAKKRPIVAHPSPMR	300
D6	241	EKSHPSVPYPHYPEKGRLDECOMYLAEQA PRSAHRIPTTEKAVFSRNAKKRP IVAHPSPMR	300
Qy	301	TE 302	
D6	301	TE 302	

RESULT 4

```

US-09-895-298-120
: Sequence 120, Application US/09895298
: Publication No. US20030078405A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: 47 Human Secreted Proteins
: FILE REFERENCE: P2035P1
: CURRENT APPLICATION NUMBER: US/09/895,298
: PRIOR FILING DATE: 2001-07-02
: PRIOR APPLICATION NUMBER: 09/591,16
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: PCT/US99/29950
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: 60/113,006
: PRIOR FILING DATE: 1998-12-18
: PRIOR APPLICATION NUMBER: 60/112,809
: PRIOR FILING DATE: 1998-12-17
: NUMBER OF SEQ ID NOS: 231
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 120
: LENGTH: 345
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (280)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-895-298-120

```

Query Match	99.8%	Score 1622	DB 10	Length 345
Best Local Similarity	99.7%	Pred. No. 7.5e-160		
Matches 301; Conservative	0	Mismatches 1	Indels 0	Gaps 0

Qy	1	MAAPGRVLVLIICSVVFSAVYIILCCMAGLPLCLATCLDHHPTGSPPTVGPPLHPSGYS	60
Db	44	MAAPGRVLIIICSVFSAVYIILCCMAGLPLCLATCLDHHPTGSRPTVGPPLHPSGYS	103
Qy	61	SVDPCKPLVREPCRSCAVYSSSGOMLGSGIABEIDSAECVFRNNOAPTVGFADVGQRST	120
Db	104	SVDPCKPLVREPCRSCAVYSSSGOMLGSGIGABEIDSAECVFRNNOAPTVGFADVGQRST	163
Qy	121	LRVSHTSVPLLRNYSHPQKARDPLIYMWGGGRHMDRLVGGRTYRITLLQLTRMPYGLQ	180
Db	164	LRVSHTSVPLLRNYSHPQKARDPLIYMWGGGRHMDRLVGGRTYRITLLQLTRMPYGLQ	223
Qy	161	VYTFERNMAVYCDQIFODETGKRRSGSFISGTWFTMLATELCEBIVVYGNVSDSYCR	240
Db	224	VYTFERNMAVYCDQIFODETGKRRSGSFISGTWFTMLATELCEBIVVYGNVSDSYCR	283
Qy	241	EKSHSVPPIHYEKEGRLDSCOMYLAEQOAPRSARHRTTEKAVSRMAKKRPIVFAHPSWR	300
Db	284	EKSHSVPPIHYEKEGRLDSCOMYLAEQOAPRSARHRTTEKAVSRMAKKRPIVFAHPSWR	343
Qy	301	TE 302	
Db	344	TE 345	

RESULT 5

US-10-262-511-184
; Sequence 184, Application US/10262511
; Publication No. US20040038223A1
; GENERAL INFORMATION:

APPLICANT: Smithson, Giendda
 APPLICANT: Millet, Isabelle
 APPLICANT: Peyman, John A.
 APPLICANT: Kekuda, Ramesh
 APPLICANT: Ju, JingTang
 APPLICANT: Li, Li
 APPLICANT: Guo, Xiaojia (Sasha)
 APPLICANT: Patturajan, Meera
 APPLICANT: Spytek, Kimberly A.
 APPLICANT: Edinger, Shlomit R.
 APPLICANT: Ellerman, Karen
 APPLICANT: Malyanekar, Uriel M.
 APPLICANT: Ort, Tatiana
 APPLICANT: Gorman, Linda
 APPLICANT: Zernhusen, Bryan D.
 APPLICANT: Anderson, David W.
 APPLICANT: Zhong, Mei
 APPLICANT: Caterton, Elina
 APPLICANT: Ji, Weizhen
 APPLICANT: Miller, Charles E.
 APPLICANT: Rastelli, Luca
 APPLICANT: Stone, David J.
 APPLICANT: Pena, Carol E. A.
 APPLICANT: Shenoy, Suresh G.
 APPLICANT: Shinkete, Richard A.
 APPLICANT: Rothenberg, Mark E.
 APPLICANT: Leach, Martin D.
 APPLICANT: Agee, Michele L.
 APPLICANT: Berghs, Constance
 TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

CURRENT APPLICATION NUMBER: US/10/262,511

```

CURRENT FILING DATE: 2003-05-28
PRIOR APPLICATION NUMBER: 60/326,483
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: 60/373,815
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/327,917
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/381,642
PRIOR FILING DATE: 2002-05-17
PRIOR APPLICATION NUMBER: 60/328,029
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: 60/381,038
PRIOR FILING DATE: 2002-05-16
PRIOR APPLICATION NUMBER: 60/328,056
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/373,260
PRIOR FILING DATE: 2002-04-17
PRIOR APPLICATION NUMBER: 60/373,826
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/327,435
PRIOR FILING DATE: 2001-10-05
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 439
SOFTWARE: Curaseq1st version 0.1
SEQ ID NO 184
LENGTH: 218
TYPE: PRT
ORGANISM: Homo sapiens
US-10-262-511-184

```

Query Match	71.8%	Score 1167;	DB 15;	Length 218;
Best Local Similarly	100.0%	Pred. No. 9.2e-11;		
Matches 218; Conservative	0;	Mismatches	0;	Gaps 0

0y 85 MMSGGLGAIIDSAECVFRNNQAPVGVFEADVGGSTLRVVSHTSVPLLNYSHYFOKAR 144

Db 1 MMSGGLGAIIDSAECVFRNNQAPVGVFEADVGGSTLRVVSHTSVPLLNYSHYFOKAR 60

Qy	145	DLTYMVGQSRHNDRIVLGSGSTRITLLQLFRMYPGQVYTFERMMAYCDQJFODETGKNR	204
Db	61	DLTYMVGQSRHNDRIVLGSGSTRITLLQLFRMPGQVYTFERMMAYCDQJFODETGKNR	120
Qy	205	ROSGSFJSTGWFMTIATLCEBIIVVYGWVSDYCREKSHSPVPHYFEKRLDECOMYL	264
Db	121	ROSGSFJSTGWFMTIATLCEBIIVVYGWVSDYCREKSHSPVPHYFEKRLDECOMYL	180
Qy	265	AHEQAPRSARHPIITEKAVFSRNAKKRPVFAHPSMRT	302
Db	181	AHEQAPRSARHPIITEKAVFSRNAKKRPVFAHPSMRT	218

```

RESULT 6
US-10-844-874-20
; Sequence 20, Application US/10844874
; Publication No. US20040204381A1
GENERAL INFORMATION:
APPLICANT: Moskal, Joseph
APPLICANT: Yamamoto, Hirotaka
; TITLE OF INVENTION: Detection and Treatment of Glyco-Enzyme-Related Diseases
; FILE REFERENCE: 97-186-E
CURRENT APPLICATION NUMBER: US/10/844,874
CURRENT FILING DATE: 2004-05-13
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 20
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Rat
FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: rat alpha 2, 6-ST
US-10-844-874-20

```

	Query Match Similarity	44.7%;	Score 726.51;	DB 17;	Length 305;	
	Best Local Similarity	48.7%;	Pred. No. 1.1e-66;			
	Matches	146;	Conservative	46;	Mismatches	85;
					Indels	23;
					Gaps	5
QY	14	SVVSASVYIILLCMAGLPLCLATCTLDHFPPT-----GSRFVYPGVLHNS-----	QY	59		
DB	12	AVSFIALCITLL-----LAMRLANDVTFPRLNLNCFGPQKTKWLPSTLLRPLQTHYXG		64		
QY	60	SSVVDGKLYLVEPCRSACAVASSGQMLTSGLAIEIDSAECVFRNMQAFLVGEADVQGRS		119		
DB	65	INVTGQEPFL-QLNCHCAVAVNSGQWGVQKVEIDRASCIRNMNNAFLTGEEDVGVWT		123		
QY	120	TLRVYSHSVVLLRLNRYSHYFOKADTLVYMWGGGRHNDRLVLTGCTTYRTLDTMTYRGL		179		
DB	124	MVRVAVSHSVVPLLNLNDPDTFFFEASTTLYLVLMGFRNMKDSQNGVYNNLTKTVDAYDPA		183		
QY	180	QVYVTFTERMAVYCDQIFQDETQKNNRQSGSPLSTQMFTMLALBELCEBIVYVGNWSDSYC		239		
DB	184	QIVYVTEQRMTYCDQVFDQETGSKEDVGSGSYLSTQMFTPLIAMDACSIIHYVGMINTETYC		243		
QY	240	REKSHSPVYVYVEFGGRGLDECOMYLAHQAPRSARFLTTEKRVAFSSMKKPRIVYAHNSW		299		
DB	244	TTBGRKRVYHYVEQK-DECNBYLLHSHAPGGRFLTTEKRVAFSSMKKGRIVVTHDPW		302		

RESULT 7
 US-09-833-245-1914
 Sequence 1914, Application US/09833245
 Publication No. US20040010134A1
 GENERAL INFORMATION:
 APPLICANT: Human Genome Sciences, Inc.
 TITLE OF INVENTION: Albumin Fusion Protein
 FILE REFERENCE: PFS46PCT
 CURRENT APPLICATION NUMBER: US/09/833,245
 CURRENT FILING DATE: 2001-04-12
 PRIOR APPLICATION NUMBER: 60/229, 358
 PRIOR FILING DATE: 2000-04-12

```

      ; PRIOR APPLICATION NUMBER: 60/256, 931
      ; PRIOR FILING DATE: 2000-12-21
      ; PRIOR APPLICATION NUMBER: 60/199, 384
      ; PRIOR FILING DATE: 2000-04-25
      ; NUMBER OF SEQ ID NOS: 2267
      ; SOFTWARE: PatentIn Ver. 2.1
      ; SEQ ID NO 1914
      ; LENGTH: 305
      ; TYPE: PRT
      ; ORGANISM: Homo sapiens
      US-09-833-245-1914

```

	Query Match	44.1%	Score 716.5	DB 11	Length 305
	Best Local Similarity	52.7%	Pred. No. 1.2e-65		
	Matches	129	Conservative	48	Mismatches 65; Indels 3; Gaps 3
QY	55	HFGSGVSSDPDGKPLVRERPCRCANVSSSGMGLGSLGAELIDSAEVEFFMNQAFTYGFEAD	114		
Db	61	HY-GIYNKKIQSPQLD-CDICALVINSNGQVAGQVGNEDIRSCSTWMNNAPTYGEED	118		
QY	115	VGORSTLRVVSHTSVPDLLRNYSYHFOKARDTLVMWOGGRHMDEVLOGRTYRTLQLTR	174		
Db	119	VGRMTMLRVVSHTSVPLLKNPDYFKEAKNTTIVIMGPFRMRKDGGIYYNMLKKTIVG	178		
QY	175	MYRGLQYITTFERRMAAYCDQIFODETGKNRQSSFLSTGMFTMLLALEBEIYVGMV	234		
Db	179	IYPAAQIVVTTEKMSYCDGVFKKETGDRVOSSYSTGWFTFLIADACGIHVAMT	238		
QY	235	SDSYCREKSHPSVEYHYEFEGRLDECCQMYLAEQAPRSARHTITEKAFFSMAKKRPIVF	294		
Db	239	NDLYCKMEGYAKVPHYHYEQGR-DECDEYFLAHNAIPYGNRPITEKKVFAKMAKKGRILF	297		
QY	295	AHPSW 299			
Db	298	THPNW 302			

```

RESULT 8
US-09-833-245-1915
Sequence 1915, Application US/09833245
Publication No. US20040010134A1
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: US/09/833,245
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/2229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1915
LENGTH: 305
TYPE: PRT
ORGANISM: Homo sapiens
US-09-833-245-1915

```

[illegible]

```
QY 175 MYPGLOYTTERMMAYCDQIFODETGKNRROSGSFLSTGWTMLALELCEBIVVYGMV 234
DB 179 IYPAQIYVTEKMSYCDGFKETGDRVQSSYSTIGMTFLILMDACYGHIYGM 238
QY 235 SDSYCREKSHPSVYHYFEKGRLEDCQMYLAHQAPRSARHPTTEKAVFSRMAKRPV 294
DB 239 NDTYCKTEGVRKRVYHYEQR-DECEYFLHEHAPYGHRFITEKKVFAWAKGHRILF 297
QY 295 AHPBW 299
DB 298 THPMW 302

RESULT 9
US-10-188-186-22
; Sequence 22, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT FILING DATE: 2002-07-02
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 22
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-186-22

Query Match 43.9%; Score 714.5; DB 15; Length 305;
Best Local Similarity 52.2%; Pred. No. 1.9e-65;
Matches 128; Conservative 49; Mismatches 65; Indels 3; Gaps 3;

QY 55 HFGSYSSVPDGKPLVREPCRSQAVSSSGQMLSGLAGAIDSAECVFRMNAQPTVGF 114
DB 61 HY-GYINVKTOEPLQD-CDLCAIVNSGQVGVQVGENIDRSSCIWRMNNAPTKGYEED 118
QY 115 VGRMTMIRVSHSVPLLNKPNDFFEKANTTYIVWGPFRNMKDGNGIYNNMLKKT 178
DB 119 VGRMTMIRVSHSVPLLNKPNDFFEKANTTYIVWGPFRNMKDGNGIYNNMLKKT 178
QY 175 MYPGLOYTTERMMAYCDQIFODETGKNRROSGSFLSTGWTMLALELCEBIVVYGMV 234
DB 179 IYPAQIYVTEKMSYCDGFKETGDRVQSSYSTIGMTFLILMDACYGHIYGM 238
QY 235 SDSYCREKSHPSVYHYFEKGRLEDCQMYLAHQAPRSARHPTTEKAVFSRMAKRPV 294
DB 239 NDTYCKTEGVRKRVYHYEQR-DECEYFLHEHAPYGHRFITEKKVFAWAKGHRILF 297
QY 295 AHPBW 299
```

```
DB 298 THPMW 302

RESULT 10
US-10-104-047-2350
; Sequence 2350, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2350
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2350

Query Match 43.9%; Score 713.5; DB 14; Length 305;
Best Local Similarity 52.2%; Pred. No. 2.4e-65;
Matches 128; Conservative 49; Mismatches 65; Indels 3; Gaps 3;

QY 55 HFGSYSSVPDGKPLVREPCRSQAVSSSGQMLSGLAGAIDSAECVFRMNAQPTVGF 114
DB 61 HY-GYINVKTOEPLQD-CDLCAIVNSGQVGVQVGENIDRSSCIWRMNNAPTKGYEED 118
QY 115 VGRMTMIRVSHSVPLLNKPNDFFEKANTTYIVWGPFRNMKDGNGIYNNMLKKT 178
DB 119 VGRMTMIRVSHSVPLLNKPNDFFEKANTTYIVWGPFRNMKDGNGIYNNMLKKT 178
QY 175 MYPGLOYTTERMMAYCDQIFODETGKNRROSGSFLSTGWTMLALELCEBIVVYGMV 234
DB 179 IYPAQIYVTEKMSYCDGFKETGDRVQSSYSTIGMTFLILMDACYGHIYGM 238
QY 235 SDSYCREKSHPSVYHYFEKGRLEDCQMYLAHQAPRSARHPTTEKAVFSRMAKRPV 294
DB 239 NDTYCKTEGVRKRVYHYEQR-DECEYFLHEHAPYGHRFITEKKVFAWAKGHRILF 297
QY 295 AHPBW 299
DB 298 THPMW 302

RESULT 11
US-10-188-186-24
; Sequence 24, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT FILING DATE: 2002-07-02
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
```


SEQ ID NO 189
LENGTH: 336
TYPE: PRT
ORGANISM: Homo sapiens
US-10-097-340-189

Query Match 39.3%; Score 639.5; DB 14; Length 336;
Best Local Similarity 47.9%; Pred. No. 1.4e-57;
Matches 125; Conservative 44; Mismatches 89; Indels 3; Gaps 3;

QY 44 TGSRPVTP-GEIHFSGYSSVPDGGKPLVREPCRSQAVSSSGOMLGSGIAEIDSACVFR 102
DB 64 TQGRPGVPAGRPRLDGLGVADHKPL-KMCRDCAVLTSSGHLHSGSQIDQTECVIR 122
QY 103 MNOAPVGFADVGORSTLAVSHTSVPLLRNYSYFQKARDTLVWVGQGRHMDVLG 162
DB 123 MNDAPTRGVRDVGNTSLRVIHSSIORILRRHDLNVSGQTVFLFWGSSYMRDCK 182
QY 163 GRITRTLLQITRMVPGIQVYTFERRMAYCDQIFQDETGNRRQSGSFLSTGFTMTIAL 222
DB 183 GQVYNNLHLISQVLPRLKAFMITRHKMLQFDELFKQETGDKRKISNTWLSTGFTMTIAL 242
QY 223 ELCEEIVYGMVSDSCREKSHSPVPHYFEKGRLDCECOMYLAEQAPR-SARHFTTEKA 281
DB 243 ELCDRIIVYGMVPPDFCRDNPSPVPHYFEPPDECTMYLSHERGKSHRRFTTEKR 302
QY 282 VFSRMKKRPVFAHPSWRT 302
DB 303 VFKMARTFNIHFQPDWKPE 323

RESULT 14

US-10-094-749-2441
Sequence 2441, Application US/10094749
Publication No. US20030219741A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
PRIOR FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2441
LENGTH: 336
TYPE: PRT
ORGANISM: Homo sapiens
US-10-094-749-2441

Query Match 39.3%; Score 639.5; DB 14; Length 336;
Best Local Similarity 47.9%; Pred. No. 1.4e-57;
Matches 125; Conservative 44; Mismatches 89; Indels 3; Gaps 3;

QY 44 TGSRPVTP-GEIHFSGYSSVPDGGKPLVREPCRSQAVSSSGOMLGSGIAEIDSACVFR 102
DB 64 TQGRPGVPAGRPRLDGLGVADHKPL-KMCRDCAVLTSSGHLHSGSQIDQTECVIR 122
QY 103 MNOAPVGFADVGORSTLAVSHTSVPLLRNYSYFQKARDTLVWVGQGRHMDVLG 162
DB 123 MNDAPTRGVRDVGNTSLRVIHSSIORILRRHDLNVSGQTVFLFWGSSYMRDCK 182
QY 163 GRITRTLLQITRMVPGIQVYTFERRMAYCDQIFQDETGNRRQSGSFLSTGFTMTIAL 222
DB 183 GQVYNNLHLISQVLPRLKAFMITRHKMLQFDELFKQETGDKRKISNTWLSTGFTMTIAL 242
QY 223 ELCEEIVYGMVSDSCREKSHSPVPHYFEKGRLDCECOMYLAEQAPR-SARHFTTEKA 281
DB 243 ELCDRIIVYGMVPPDFCRDNPSPVPHYFEPPDECTMYLSHERGKSHRRFTTEKR 302
QY 282 VFSRMKKRPVFAHPSWRT 302
DB 303 VFKMARTFNIHFQPDWKPE 323

RESULT 15

US-10-288-252-11
Sequence 11, Application US/10288252
Publication No. US20030143686A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: LAL, Preeti G.
APPLICANT: TANG, Y. Tom
APPLICANT: YUE, Henry
APPLICANT: BURFORD, Neil
APPLICANT: GANDHI, Ameena R.
APPLICANT: WARREN, Bridget A.
APPLICANT: YAO, Monique G.
APPLICANT: TRIBOULEY, Catherine M.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: LEE, Ernestine A.
APPLICANT: HAPALIA, April J.A.
APPLICANT: LU, Yan
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: SANJANMALA, Madhu S.
APPLICANT: DING, Li
TITLE OF INVENTION: TRANSFERASES
FILE REFERENCE: PI-0241 USA
CURRENT APPLICATION NUMBER: US/10/288,252
PRIOR FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: PCT US01/30424
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US 60/252,819
PRIOR FILING DATE: 2000-11-21
PRIOR APPLICATION NUMBER: US 60/249,639
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: US 60/247,931
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/246,001
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US 60/244,025
PRIOR FILING DATE: 2000-10-27
PRIOR APPLICATION NUMBER: US 60/238,481
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/236,523
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PERL Program
SEQ ID NO 11
LENGTH: 336
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: incyte id No. US20030143686A1 278782CD1
US-10-288-252-11

Query Match	39.1%;	Score 636.5;	DB 14;	Length 336;
Best Local Similarity	47.5%;	Pred. No. 2.9e-57;		
Matches 124;	Conservative 45;	Mismatches 89;	Indels 3;	Gaps 3;

Qy	4	TGSRPTVP-GPIHFGSGSSVPDGKPLVVRPPGSCAVVSSSQMGSGGLGAIIDAEVCVR	102
	64	TQGRPGVPRGRRLDGLVGLADHKPL-KKHCDCALVYSSGHLHSHRQSGSIDDTCEVTR	122
Qy	103	MNOAPVGFGEADVGRSTLRVVSHTSVPLLRNYSHYFOKARDTLVYMGQGRMRDVLG	162
Db	123	MNAPPRGYGRDVGNTSRSLVIAHSSIORILINRDLINVSQGVFIEMGSSSYMRDGGK	182
Qy	163	GRTYRTLLQTRVRYPGILOYVFTFERMMAVCDQIFQDETGGKRRQSGSPFLSGWFTMIAL	222
Db	183	GQYNNHLHLISQVLPYLKAFMTIRKMLQFDELFOEJGGRKKSINWYLSGWFTMIAL	242
Qy	223	ELCEEIVYGVMSDYSYCREKSHSPVYHYFEKGRLLDECOMYLAHQAPR-SAHRRFITEKA	281
Db	243	ELCDRLINVGVMPDPDCCRPNNHSPVYHYDDPFGDECTMYLSHBRGKSGHHRFITEKR	302
Qy	282	VFSRMAKRPVPAHPSWRTS	302
Db	303	VFKQMAFTPIHFQPDWKAPE	323

Search completed: December 10, 2004, 18:29:16
Job time : 148 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2004, 18:03:47 ; Search time 156 Seconds
(without alignments)
694.463 Million cell updates/sec

Title: US-10-019-114A-7
Perfect score: 1626
Sequence: 1 MKAFGRLLVILICVFESAV.....FSRWAKRPPIVFAHPSWRT 302

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729239 residues
Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1626	100.0	302 3 AAB25764	AAB25764 Human sec
2	1626	100.0	302 3 AAB28674	AAB28674 Human car
3	1626	100.0	302 4 AAB61614	AAB61614 Human pro
4	1626	100.0	302 4 AAB75350	AAB75350 Human sec
5	1626	100.0	302 7 ADD47793	ADD47793 Human pro
6	1626	100.0	302 8 ADP19151	ADP19151 Human sec
7	1622	99.8	303 3 AAB24441	AAB24441 Human sec
8	1622	99.8	345 3 AAB24495	AAB24495 Human sec
9	1167	71.8	218 6 ADA05824	ADA05824 Human NOV
10	1159	71.3	218 6 ADN62987	ADN62987 Human NOV
11	726.5	44.7	305 7 ADD47791	ADD47791 Rat Prote
12	716.5	44.1	305 5 AAE21578	AAE21578 Human gen
13	716.5	44.1	305 5 AAE21595	AAE21595 Human gen
14	716.5	44.1	305 5 ABG65165	ABG65165 Human alb
15	716.5	44.1	305 5 ABG65166	ABG65166 Human alb
16	716.5	44.1	305 8 ADL78432	ADL78432 Albumin f
17	716.5	44.1	305 8 ADL78433	ADL78433 Albumin f
18	716.5	44.1	305 8 AAE21617	AAE21617 Human gen
19	714.5	43.9	305 7 ADC26197	ADC26197 Human NOV
20	713.5	43.9	305 7 ADB64196	ADB64196 Human pro
21	694	42.7	318 7 ADC26199	ADC26199 Human NOV
22	677.5	41.7	305 7 AAE10705	AAE10705 Human NOV
23	639.5	39.3	336 4 AAE10705	AAE10705 Human s1a
24	639.5	39.3	336 5 ABG96367	ABG96367 Human ova
25	639.5	39.3	336 6 ADA54873	ADA54873 Human pro

26	636.5	39.1	336 5 AAE22149	AAE22149 Human TRN
27	633.5	39.0	275 4 AAM41882	AAM41882 Human pol
28	595.5	36.6	246 4 AAM40096	AAM40096 Human pol
29	548.5	33.7	256 4 AAM41144	AAM41144 Human pol
30	537	33.0	299 3 AAY99356	AAY99356 Human PRO
31	537	33.0	299 3 AAB66105	AAB66105 Protein o
32	537	33.0	299 4 AAU29126	AAU29126 Human PRO
33	537	33.0	299 6 ABUS5802	ABUS5802 Human PRO
34	537	33.0	299 6 ABUS8050	ABUS8050 Novel hum
35	537	33.0	299 6 ABUS8365	ABUS8365 Human sec
36	537	33.0	299 6 ABR66239	ABR66239 Human sec
37	537	33.0	299 6 ABR65629	ABR65629 Human sec
38	537	33.0	299 6 ABUS9569	ABUS9569 Human sec
39	537	33.0	299 6 ABUS8208	ABUS8208 Human PRO
40	537	33.0	299 6 ABUS9929	ABUS9929 Novel hum
41	537	33.0	299 6 ABR68178	ABR68178 Human sec
42	537	33.0	299 6 ABUS96231	ABUS96231 Novel hum
43	537	33.0	299 6 ABUS92662	ABUS92662 Human sec
44	537	33.0	299 6 ABO08739	ABO08739 Human sec
45	537	33.0	299 6 ABO02791	ABO02791 Human sec

ALIGNMENTS

RESULT 1
AAB25764
ID AAB25764 standard; protein; 302 AA.
XX
XX AAB25764;
DT 28-NOV-2000 (first entry) date of database entry

DE Human secreted protein SEQ ID #76.

XX
XX Human; secreted protein; forensic procedure; gene therapy;
KW Chromosome mapping; cancer; autoimmune disease; cardiovascular disorder;
KW Cystic fibrosis; hypothyroidism; immunological disorder; amyloidosis;
KW brain disorder; skeletal muscle disorder; eye disorder; obesity;
KW mitochondrial cytopathy; diabetes; atherosclerosis; Alzheimer's disease;
KW neurodegenerative disorder; graft rejection; dementia; hyperlipidaemia;
KW septic shock; impotence.

XX Homo sapiens.

XX WO200037491-A2.

PD 29-JUN-2000.

PR 20-DEC-1999; 99NC-IB002058.

PR 25-JUN-1999; 98US-0113686P.

XX (GERT) GENSET.

XX Bougueleret L, Dumas J, Duclert A;

XX WPI; 2000-442637/38.

DR N-PSDB; AAA87726.

XX Polynucleotides and polypeptides encoding proteins with signal peptides,

PT useful in diagnostic, forensic, gene therapy and chromosome mapping

PS Claim 9; Fig 12; 306pp; English.

XX This sequence represents a human secreted protein amino acid sequence.
CC The invention relates to sequences AAB25763-B25812. The proteins include signal peptides.
CC secreted proteins AAB25763-B25812. The proteins include one of the CDNA
CC sequences, and a purified antibody capable of binding to one of the
CC secreted proteins. Also contained in the invention are methods for

Patent publication date / issued
Patent filing date / submitted
Priority date

CC storing the sequence data on a computer system, and a method for
 CC identifying features of the cDNA sequences using a computer programme.
 CC The cDNAs are useful for expressing secreted proteins or fragments to
 CC obtain antibodies capable of specifically binding to the secreted
 CC proteins. The cDNAs may also be useful in diagnostic, forensic, gene
 CC therapy and chromosome mapping procedures and may be used to design
 CC expression vectors and secretion vectors. The proteins of the invention
 CC may be used to treat diseases including cancer, autoimmune diseases,
 CC cardiovascular disorders, cystic fibrosis, hypothyroidism, immunological
 CC disorders, amyloidosis, brain disorders, skeletal muscle disorders, eye
 CC disorders, obesity, mitochondrial cytopathies, diabetes, atherosclerosis,
 CC neurodegenerative disorders, graft rejection, Alzheimer's disease,
 CC dementia, hyperlipidaemia, septic shock and impotence
 CC
 XX
 XX Sequence 302 AA;
 SQ
 Query Match 100.0%; Score 1626; DB 3; Length 302;
 Best Local Similarity 100.0%; Pred. No. 1.2e-160;
 Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKAPGRVLVILICSVFSAVYIILCCWAGLPCLCLATCLDHHFPTGSRPTVGPGLHFSGYG 60
 DB 1 MKAPGRVLVILICSVFSAVYIILCCWAGLPCLCLATCLDHHFPTGSRPTVGPGLHFSGYG 60
 QY 61 SVPDGKPLVREPCRSQAVVSSSGQMLGSGIAGIIDSACVFRNMQAPTVGFENDVGQRST 120
 DB 61 SVPDGKPLVREPCRSQAVVSSSGQMLGSGIAGIIDSACVFRNMQAPTVGFENDVGQRST 120
 QY 121 LRVVSHTSVPLLRNYSHFQKARDTLVYWVGGRHMDRVLGGRYRTLLQLTRMPGLQ 180
 DB 121 LRVVSHTSVPLLRNYSHFQKARDTLVYWVGGRHMDRVLGGRYRTLLQLTRMPGLQ 180
 QY 181 VYFTTERMAVYCDQIFQDETGNKRRSGSFLSTGFTMTIALBELCEIIVVYGWVSDSYCR 240
 DB 181 VYFTTERMAVYCDQIFQDETGNKRRSGSFLSTGFTMTIALBELCEIIVVYGWVSDSYCR 240
 QY 241 EKSHPSVPHYFEKGRLDCEOMYLAHQAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300
 DB 241 EKSHPSVPHYFEKGRLDCEOMYLAHQAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300
 QY 301 TE 302
 DB 301 TE 302
 RESULT 2
 AAB28674
 ID AAB28674 standard; protein; 302 AA.
 AC AAB28674;
 XX
 DT 13-FEB-2001 (first entry)
 DE Human carbohydrate-modifying enzyme Incyte ID No: 983984CD1.
 XX
 KW Human; carbohydrate-modifying enzyme; CME; antidiabetic;
 KW immunosuppressive; anti-HIV; antiinflammatory; antianaemic;
 KW antidiabetic; antiarteriosclerotic; antihypertensive; hepatotropic;
 KW nephrotropic; antitumor; thyromimetic; neuroprotective; osteoparitic;
 KW antidiabetic; antipsoriatic; uroparitic; ophthalmological;
 KW dermatological; antitumor; cytostatic; virucide; antibacterial;
 KW fungicide; protozoicide; tranquiliser; vulnerary; diabetes;
 KW autoimmune disorder; inflammatory disorder; infection.
 XX
 XX Homo sapiens.
 OS
 XX
 PN WO200063351-A2.
 XX
 PD 26-OCT-2000.
 XX
 XX 20-APR-2000; 2000WO-US010882.
 PF
 XX 21-APR-1999; 99US-013083P.
 PR

XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Lal P, Yue H, Tang YT, Hillman JL, Baughn MR, Yang J;
 XX
 XX WPI; 2000-672729/65.
 DR N-PSDB; AAC65393.
 XX
 PT Novel carbohydrate modifying enzyme polypeptides and polynucleotides for
 PT diagnosis, treatment, and prevention of carbohydrate metabolism
 PT disorders, autoimmune/inflammatory disorders, and cancer.
 PS
 PS Claim 1; Page 67-68; 75pp; English.
 XX
 CC The present sequence is a human carbohydrate-modifying enzyme (CME). CME
 CC polynucleotides and polypeptides are useful for treating and diagnosing
 CC diseases associated with CME such as diabetes, autoimmune/inflammatory
 CC disorders such as AIDS, Addison's disease, adult respiratory distress
 CC syndrome, allergies, anaemia, asthma, atherosclerosis, autoimmune
 CC thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's
 CC disease, emphysema, erythroblastosis fetalis, glomerulonephritis, Good
 CC pasture's syndrome, gout, Grave's disease, Hashimoto's thyroiditis,
 CC multiple sclerosis, myasthenia gravis, osteoarthritis, osteoporosis,
 CC pancreatitis, polymyositis, psoriasis, Reiter's syndrome, arthritis,
 CC scleroderma, Sjogren's syndrome, systemic lupus erythematosus, ulcerative
 CC colitis, uveitis, Werner syndrome, complications of cancer,
 CC haemodialysis, and extracorporeal circulation, viral, bacterial, fungal
 CC parasite, protozoal, and helminthic infections, trauma, or cancer. CME,
 CC or its catalytic or immunogenic fragment, is useful for drug screening
 CC
 XX
 SQ Sequence 302 AA;
 Query Match 100.0%; Score 1626; DB 3; Length 302;
 Best Local Similarity 100.0%; Pred. No. 1.2e-160;
 Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKAPGRVLVILICSVFSAVYIILCCWAGLPCLCLATCLDHHFPTGSRPTVGPGLHFSGYG 60
 DB 1 MKAPGRVLVILICSVFSAVYIILCCWAGLPCLCLATCLDHHFPTGSRPTVGPGLHFSGYG 60
 QY 61 SVPDGKPLVREPCRSQAVVSSSGQMLGSGIAGIIDSACVFRNMQAPTVGFENDVGQRST 120
 DB 61 SVPDGKPLVREPCRSQAVVSSSGQMLGSGIAGIIDSACVFRNMQAPTVGFENDVGQRST 120
 QY 121 LRVVSHTSVPLLRNYSHFQKARDTLVYWVGGRHMDRVLGGRYRTLLQLTRMPGLQ 180
 DB 121 LRVVSHTSVPLLRNYSHFQKARDTLVYWVGGRHMDRVLGGRYRTLLQLTRMPGLQ 180
 QY 181 VYFTTERMAVYCDQIFQDETGNKRRSGSFLSTGFTMTIALBELCEIIVVYGWVSDSYCR 240
 DB 181 VYFTTERMAVYCDQIFQDETGNKRRSGSFLSTGFTMTIALBELCEIIVVYGWVSDSYCR 240
 QY 241 EKSHPSVPHYFEKGRLDCEOMYLAHQAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300
 DB 241 EKSHPSVPHYFEKGRLDCEOMYLAHQAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300
 QY 301 TE 302
 DB 301 TE 302
 RESULT 3
 AAB61614
 ID AAB61614 standard; protein; 302 AA.
 AC AAB61614;
 XX
 DT 05-APR-2001 (first entry)
 DE Human protein HP03380.
 XX
 XX Human; hydophobic domain; immune deficiency; autoimmune disorder;
 KW allergy; tissue growth; regeneration; wound healing; burn; tumour;
 KW

KM periodontal disease; thrombolytic condition; haemostatic condition;
XX infection.
XX Homo sapiens.
XX MO200102563-A2.
XX
XX 11-JAN-2001.
XX
XX 16-JUN-2000; 2000MO-JP003943.
XX
XX 02-JUL-1999; 99JP-00188835.
XX
XX (SAGA) SAGAMI CHEM RES CENT.
XX (PROT-) PROTEGENE INC.
XX Kato S, Kimura T;
XX WPI; 2001-071581/08.
XX N-PSDB; AAF28685, AAF28695.
XX
XX New human proteins with hydrophobic domains, useful for the treatment of
XX immune disorders, tumors, allergic conditions, thrombosis and microbial
XX infection.
XX
XX Claim 1; Page 105-106; 153pp; English.
XX
XX The present invention relates to human proteins (AAB61608-AAB61617) and
XX their coding sequences (AAF28679-AAF28698). The proteins of the present
XX invention have hydrophobic domains and can be used for the treatment of
XX various immune deficiencies and disorders, such as severe combined
XX immunodeficiency (SCID), multiple sclerosis, rheumatoid arthritis,
XX autoimmune pulmonary inflammation, graft-versus-host disease and Guillain
XX-Barre syndrome. The proteins may also be useful in the treatment of
XX allergic reactions and conditions, such as asthma and in regulation of
XX hematopoiesis or lymphoid cell deficiencies. The proteins may also have
XX utility in compositions used for bone, cartilage, tendon and/or nerve
XX tissue growth or regeneration as well as wound healing and in the
XX treatment of burns. The proteins may be used in the treatment of
XX periodontal disease and in other tooth repair processes. Other uses
XX include treatment of thrombolytic and haemostatic conditions, treatment
XX or prevention of tumours and inhibiting infection by bacteria, viruses,
XX fungi and other parasites
XX
XX Sequence 302 AA;
SQ
Query Match 100.0%; Score 1626; DB 4; Length 302;
Best Local Similarity 100.0%; Pred. No. 1.2e-160;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKAPGRVLVLIILCSVFSAYIILCCWAGLPCLATCLDHHFPTGSRPTVPGPLHFGYS 60
DB 1 MKAPGRVLVLIILCSVFSAYIILCCWAGLPCLATCLDHHFPTGSRPTVPGPLHFGYS 60
QY 61 SVDPGKPLVBEPCRSQCAVSSSGOMLGGAGLIDSAECVFRMNOAFTVGFADVGQRST 120
DB 61 SVDPGKPLVBEPCRSQCAVSSSGOMLGGAGLIDSAECVFRMNOAFTVGFADVGQRST 120
QY 121 LRVVSHTSVPLLRNYSHYFOKADTLVYMWGGRHNDRLVGGRTYTTLLQLTMYGQLQ 180
DB 121 LRVVSHTSVPLLRNYSHYFOKADTLVYMWGGRHNDRLVGGRTYTTLLQLTMYGQLQ 180
QY 181 VYFTEERMAVYCDQIFODETGKNRROSGSFLSTGMFTMIILALCEBIIVYGMVSDSYCR 240
DB 181 VYFTEERMAVYCDQIFODETGKNRROSGSFLSTGMFTMIILALCEBIIVYGMVSDSYCR 240
QY 241 EKSHPSVPYHYFEKGRLECOMYLAAHQAPRSARFTTEKAVFSRMAKGRPIVFAHPSWR 300
DB 241 EKSHPSVPYHYFEKGRLECOMYLAAHQAPRSARFTTEKAVFSRMAKGRPIVFAHPSWR 300
QY 301 TE 302
DB 301 TE 302

RESULT 4
AAB75350
ID AAB75350 standard; protein; 302 AA.
XX
XX AAB75350;
XX
XX 05-APR-2001 (first entry)
XX
XX Human secreted protein #9.
XX
XX Secreted protein; prevention; treatment; diagnosis; disease; infection.
XX
XX Homo sapiens.
XX
XX MO200100806-A2.
XX
XX 04-JAN-2001.
XX
XX 21-JUN-2000; 2000MO-IB000951.
XX
XX 25-JUN-1999; 99US-0141032P.
XX
XX 21-DEC-1999; 99US-00469099.
XX
XX (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Bougueleret L, Joubert S;
XX WPI; 2001-071487/08.
XX
XX 49 Secreted proteins and the nucleic acids encoding them, useful in gene
XX therapy and for detecting similar sequences in samples.
XX
XX Claim 10; Page 280; 307pp; English.
XX
XX The present invention relates to 49 Secreted proteins and the cDNAs
XX encoding them. The protein and nucleic acids may be used in the
XX prevention, treatment and diagnosis of diseases associated with
XX inappropriate protein expression
XX
XX Sequence 302 AA;
SQ
Query Match 100.0%; Score 1626; DB 4; Length 302;
Best Local Similarity 100.0%; Pred. No. 1.2e-160;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKAPGRVLVLIILCSVFSAYIILCCWAGLPCLATCLDHHFPTGSRPTVPGPLHFGYS 60
DB 1 MKAPGRVLVLIILCSVFSAYIILCCWAGLPCLATCLDHHFPTGSRPTVPGPLHFGYS 60
QY 61 SVDPGKPLVBEPCRSQCAVSSSGOMLGGAGLIDSAECVFRMNOAFTVGFADVGQRST 120
DB 61 SVDPGKPLVBEPCRSQCAVSSSGOMLGGAGLIDSAECVFRMNOAFTVGFADVGQRST 120
QY 121 LRVVSHTSVPLLRNYSHYFOKADTLVYMWGGRHNDRLVGGRTYTTLLQLTMYGQLQ 180
DB 121 LRVVSHTSVPLLRNYSHYFOKADTLVYMWGGRHNDRLVGGRTYTTLLQLTMYGQLQ 180
QY 181 VYFTEERMAVYCDQIFODETGKNRROSGSFLSTGMFTMIILALCEBIIVYGMVSDSYCR 240
DB 181 VYFTEERMAVYCDQIFODETGKNRROSGSFLSTGMFTMIILALCEBIIVYGMVSDSYCR 240
QY 241 EKSHPSVPYHYFEKGRLECOMYLAAHQAPRSARFTTEKAVFSRMAKGRPIVFAHPSWR 300
DB 241 EKSHPSVPYHYFEKGRLECOMYLAAHQAPRSARFTTEKAVFSRMAKGRPIVFAHPSWR 300
QY 301 TE 302
DB 301 TE 302

RESULT 5

ADD47793
 ID ADD47793 standard; protein; 302 AA.
 AC ADD47793;
 DT 29-JAN-2004 (first entry)
 DE Human Protein CAC07404, SEQ ID NO 13489.
 XX Human; pain; neuronal tissue; gene therapy;
 KM spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX Homo sapiens.
 OS
 PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002MO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GENO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 PI Woolf C, D'urso D, Befort K, Costigan M;
 DR WPI; 2003-268312/26.
 DR GENBANK; CAC07404.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 preparing a medicament for treating pain in an animal.
 XX
 Claim 1; Page; 1017p; English.
 XX
 The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 302 AA;

Query Match 100.0%; Score 1626; DB 7; Length 302;
 Best Local Similarity 100.0%; Pred. No. 1,2e-160;
 Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAPGRVLVLIILCSVFSAVYIILCCWAGLPCLATCLDHHFPTGSRPTVGPGLHFSGYS 60

Db 1 MKAPGRVLVLIILCSVFSAVYIILCCWAGLPCLATCLDHHFPTGSRPTVGPGLHFSGYS 60
 QY 61 SVPDGKPLVREPERSCAVYSSSGOMGSGIGATIDAECEFRNNOAPTVGEADVGQRST 120
 Db 61 SVPDGKPLVREPERSCAVYSSSGOMGSGIGATIDAECEFRNNOAPTVGEADVGQRST 120
 QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRYLGGRTYRTLLQLTRMYPGLQ 180
 Db 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRYLGGRTYRTLLQLTRMYPGLQ 180
 QY 181 VYTFTERMAVYCDQIFODETGKRRSGSFLSTGFTMIALALCEIIVYGVNVSYSYCR 240
 Db 181 VYTFTERMAVYCDQIFODETGKRRSGSFLSTGFTMIALALCEIIVYGVNVSYSYCR 240
 QY 241 EKSHPSVPHYFEKGRGLDECOMTLAHQAPRSARHPTTEKAVFSRNKKRPVFAHPSWR 300
 Db 241 EKSHPSVPHYFEKGRGLDECOMTLAHQAPRSARHPTTEKAVFSRNKKRPVFAHPSWR 300
 QY 301 TE 302
 Db 301 TE 302
 QY 301 TE 302
 Db 301 TE 302
 RESULT 6
 ADP19151
 ID ADP19151 standard; protein; 302 AA.
 XX
 AC ADP19151;
 XX
 DT 26-AUG-2004 (first entry)
 DT
 XX
 DE Human secreted polypeptide #2.
 XX
 KW Human; secreted protein; genetic disease.
 OS
 OS Homo sapiens.
 XX
 PN US2004110939-A1.
 PD 10-JUN-2004.
 XX
 PF 15-OCT-2001; 2001US-00978360.
 XX
 PR 17-DEC-1998; 98MO-IB002122.
 PR 09-FEB-1999; 99MO-IB000282.
 PR 21-JUN-2000; 2000MO-IB000951.
 PR 15-SEP-2000; 2000US-00663600.
 XX
 PA (GEST) GENSET SA.
 XX
 PI Dumas MLine Edwards J, Bougueleret L, Joberet S, Clusel C;
 PI Duclert A;
 XX
 DR WPI; 2004-440404/41.
 DR N-PSDB; ADP18746.
 XX
 PT New isolated polynucleotide encoding secreted polypeptide, useful for
 PT gene therapy, or in diagnostic procedures to identify individuals having
 PT genetic diseases resulting from abnormal expression of the genes.
 XX
 PS Claim 2; SEQ ID NO 407; 113p; English.
 XX
 The invention relates to human cDNA sequences that encode human secreted
 CC proteins. The invention also relates to an antibody that specifically
 CC binds to a polypeptide of the invention and a method of binding the
 CC polypeptide to an antibody. The polynucleotides are useful for expressing
 CC the entire secreted proteins which they encode and for distinguishing
 CC human tissues and cells from non-human tissues and cells, and for
 CC distinguishing between human tissues and cells that do or do not express
 CC the polynucleotides comprising the cDNAs. The polynucleotides and
 CC polypeptides are useful in forensic procedures or diagnostic procedures
 CC to identify individuals with genetic diseases resulting from abnormal

expression of the genes corresponding to the cDNAs. The sequences are also useful in gene therapy to control or treat genetic diseases. This sequence represents a human secreted polypeptide of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 302 AA:

Query Match 100.0%; Score 1626; DB 8; Length 302;
Best Local Similarity 100.0%; Pred. No. 1.2e-160;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAPGRVLVILLCVSVFSAVYIILCCWAGPLCLATCTLDHHPRTGSRPTVGPLHFGYS 60
DB 1 MKAPGRVLVILLCVSVFSAVYIILCCWAGPLCLATCTLDHHPRTGSRPTVGPLHFGYS 60
QY 61 SVPDGRPLVREPCRSQAVSSSGOMLGGAGETIDSAECVFRMNOAPTGVFEADVGRST 120
DB 61 SVPDGRPLVREPCRSQAVSSSGOMLGGAGETIDSAECVFRMNOAPTGVFEADVGRST 120
QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
QY 181 VYFTTERMAVYCDQIFODETGKRRROSGSLSTGFMITALECEIIVVYGWVSDSYCR 240
DB 181 VYFTTERMAVYCDQIFODETGKRRROSGSLSTGFMITALECEIIVVYGWVSDSYCR 240
QY 241 EKSHPSVPHYFEKGRLEDECOMYLAEHQAPRSARHRTTEKAVFSRMKKRPVFAHPSMR 300
DB 241 EKSHPSVPHYFEKGRLEDECOMYLAEHQAPRSARHRTTEKAVFSRMKKRPVFAHPSMR 300
QY 301 TE 302
DB 301 TE 302

RESULT 7
AAB24441
ID AAB24441 standard; protein; 303 AA.

XX AAB24441;

XX 20-NOV-2000 (first entry)

XX Human secreted protein sequence encoded by gene 5 SEQ ID NO:166.

XX Human; secreted protein; cytosolic; antianaemic; antidiabetic;
XX antiinflammatory; ophthalmological; antirheumatic; antiaortic;
XX antiisoriatic; antiangiogenic; cardiant; anti-HIV; nootropic;
XX neuroprotective; antimicrobial; antiparkinsonian; cancer;
XX immune system disorder; angiogenesis; hyperproliferative disorder;
XX cardiovascular disorder; apoptosis; neurological disease;
XX infectious disease; wound healing; chromosome 9.

XX Homo sapiens.

XX MO200035937-A1.

XX 22-JUN-2000.

XX 16-DEC-1999; 99WO-US029950.

XX 17-DEC-1998; 98US-0112809P.

XX 18-DEC-1998; 98US-0113006P.

XX (HUMA-) HUMAN GENOME SCT INC.

XX Ruben SM, Ebner R, Rosen CA, Endress GA, Soppet DR, Ni J,
XX Duan DR, Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence K,
XX WPI; 2000-431566/37.

DR N-PSDB; AAA78385.

XX Forty seven human nucleic acids encoding secreted proteins, useful in the
XX treatment, prevention and diagnosis of cancers, disorders of the immune
XX system, angiogenesis disorders, neurological diseases and
XX hyperproliferative disorders.

XX Claim 11; Page 482-483; 562pp; English.

XX The polynucleotide sequence given in AAA78381 to AAA78432 encode the
XX human secreted proteins given in AAB24437 to AAB24604. Human secreted
XX proteins have activities based on the tissues and cells the genes are
XX expressed in. Examples of activities include: cytosolic; antianaemic;
XX antidiabetic; antiinflammatory; ophthalmological; antirheumatic;
XX antiaortic; antiparkinsonian; antidiabetic; cardiant; anti-HIV;
XX nootropic; neuroprotective; antimicrobial; antiparkinsonian. Human
XX secreted protein polynucleotides, polypeptides, and/or diagnosing other
XX agonists may be useful in treating, preventing, and/or diagnosing other
XX diseases, disorders, and/or conditions such as: (a) cancers; (b)
XX disorders of the immune system; (c) angiogenesis disorders; (d)
XX hyperproliferative disorders; (e) cardiovascular disorders; (f) diseases
XX associated with increase apoptosis; (g) neurological diseases; and (h)
XX infectious diseases. They are also used to promote wound healing.
XX AAA78372 to AAA78380 and AAB24436 represent sequences used in the
XX exemplification of the present invention

XX Sequence 303 AA:

Query Match 99.8%; Score 1622; DB 3; Length 303;
Best Local Similarity 99.7%; Pred. No. 3.3e-160;
Matches 301; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKAPGRVLVILLCVSVFSAVYIILCCWAGPLCLATCTLDHHPRTGSRPTVGPLHFGYS 60
DB 1 MKAPGRVLVILLCVSVFSAVYIILCCWAGPLCLATCTLDHHPRTGSRPTVGPLHFGYS 60
QY 61 SVPDGRPLVREPCRSQAVSSSGOMLGGAGETIDSAECVFRMNOAPTGVFEADVGRST 120
DB 61 SVPDGRPLVREPCRSQAVSSSGOMLGGAGETIDSAECVFRMNOAPTGVFEADVGRST 120
QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
QY 181 VYFTTERMAVYCDQIFODETGKRRROSGSLSTGFMITALECEIIVVYGWVSDSYCR 240
DB 181 VYFTTERMAVYCDQIFODETGKRRROSGSLSTGFMITALECEIIVVYGWVSDSYCR 240
QY 241 EKSHPSVPHYFEKGRLEDECOMYLAEHQAPRSARHRTTEKAVFSRMKKRPVFAHPSMR 300
DB 241 EKSHPSVPHYFEKGRLEDECOMYLAEHQAPRSARHRTTEKAVFSRMKKRPVFAHPSMR 300
QY 301 TE 302
DB 301 TE 302

RESULT 8
AAB24495
ID AAB24495 standard; protein; 345 AA.

XX AAB24495;

XX 20-NOV-2000 (first entry)

XX Human secreted protein sequence encoded by gene 5 SEQ ID NO:120.

XX Human; secreted protein; cytosolic; antianaemic; antidiabetic;
XX antiinflammatory; ophthalmological; antirheumatic; antiaortic;
XX antiisoriatic; antiangiogenic; cardiant; anti-HIV; nootropic;
XX neuroprotective; antimicrobial; antiparkinsonian; cancer;
XX immune system disorder; angiogenesis; hyperproliferative disorder;
XX cardiovascular disorder; apoptosis; neurological disease;

KM infectious disease; wound healing; chromosome 9.
 XX Homo sapiens.
 OS
 XX WO200035937-A1.
 XX
 XX 22-JUN-2000.
 XX
 XX 16-DEC-1999; 99WO-US029950.
 PF
 XX 17-DEC-1998; 98US-0112809P.
 XX 18-DEC-1998; 98US-0113006P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ruben SM, Ebner R, Rosen CA, Endress GA, Soppet DR, Ni J,
 PI Dhan DR, Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence K,
 XX WPI; 2000-431566/37.
 DR
 XX Forty seven human nucleic acids encoding secreted proteins, useful in the
 PT treatment, prevention and diagnosis of cancers, disorders of the immune
 PT system, angiogenesis disorders, neurological diseases and
 PT hyperproliferative disorders.
 XX
 XX Disclosure; Page 19; 562pp; English.
 PS
 XX The polynucleotide sequence given in AAA78381 to AAA78432 encode the
 CC human secreted proteins given in AAB24437 to AAB24604. Human secreted
 CC proteins have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: cytostatic; antineoplastic;
 CC antidiabetic; antiinflammatory; ophthalmological; antirheumatic;
 CC antituberc; antiparasitic; antianemic; cardiatic; anti-HIV;
 CC neurotrophic; neuroprotective; antimicrobial and antiparkinsonian. Human
 CC secreted protein polynucleotides, polypeptides, antagonists and/or
 CC agonists may be useful in treating, preventing, and/or diagnosing other
 CC diseases, disorders, and/or conditions such as: (a) cancers; (b)
 CC disorders of the immune system; (c) angiogenesis disorders; (d)
 CC hyperproliferative disorders; (e) cardiovascular disorders; (f) diseases
 CC associated with increase apoptosis; (g) neurological diseases; and (h) diseases
 CC infectious diseases. They are also used to promote wound healing.
 CC AAA78372 to AAA78380 and AAB24436 represent sequences used in the
 CC exemplification of the present invention
 CC
 XX Sequence 345 AA;
 SO
 Query Match 99.8%; Score 1622; DB 3; Length 345;
 Best Local Similarity 99.7%; Pred. No. 3.9e-160;
 Matches 301; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MKAPGRVLIIICSVSAVYIIICWAGPLCLATCLDHHFPTGSRPTVPGPLHFGYS 60
 DB 44 MKAPGRVLIIICSVSAVYIIICWAGPLCLATCLDHHFPTGSRPTVPGPLHFGYS 103
 QY 61 SVDPGKLVLPFCSCAVSSSGMTGSGIAGIIDSACEVFRNQAPTVGFADVGQST 120
 DB 104 SVDPGKLVLPFCSCAVSSSGMTGSGIAGIIDSACEVFRNQAPTVGFADVGQST 163
 QY 121 LRVVSHSVPLLIANNYSHYFOKADTLVYMWGQGRHMDRVLGRTYTLIQLTRMYPGLQ 180
 DB 164 LRVVSHSVPLLIANNYSHYFOKADTLVYMWGQGRHMDRVLGRTYTLIQLTRMYPGLQ 223
 QY 181 VYFETERMAVYCDIODETGKRRRSGSFSTGFMITLALCEITVYGVWSDYCR 240
 DB 224 VYFETERMAVYCDIODETGKRRRSGSFSTGFMITLALCEITVYGVWSDYCR 283
 QY 241 EKSHPSVYHYFEKGRIDECOMYLAHQAPRSARHPTTEKAVSRMAKRPVFAHPSMR 300
 DB 284 EKSHPSVYHYFEKGRIDECOMYLAHQAPRSARHPTTEKAVSRMAKRPVFAHPSMR 343
 QY 301 TE 302
 DB 344 TE 345

RESULT 9
 ADA05824
 ID ADA05824 standard; protein; 218 AA.
 XX
 XX ADA05824;
 AC
 XX 06-NOV-2003 (first entry)
 DT
 XX Human NOV44a protein SEQ ID NO:184.
 DE
 XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
 XX immunomodulator; cytostatic; neurotrophic; neuroprotective;
 KW antiparkinsonian; antidiabetic; gene therapy; human disease;
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; hematopoietic disorder; dyslipidaemia.
 XX
 XX Homo sapiens.
 OS
 XX WO2003029424-A2.
 PN
 XX 10-APR-2003.
 PD
 XX
 XX 02-OCT-2002; 2002WO-US031373.
 PF
 XX 02-OCT-2001; 2001US-0326483P.
 PR 05-OCT-2001; 2001US-0327435P.
 PR 05-OCT-2001; 2001US-0327449P.
 PR 09-OCT-2001; 2001US-0327917P.
 PR 09-OCT-2001; 2001US-0328029P.
 PR 09-OCT-2001; 2001US-0328044P.
 PR 09-OCT-2001; 2001US-0328056P.
 PR 12-OCT-2001; 2001US-0328849P.
 PR 15-OCT-2001; 2001US-0329414P.
 PR 17-OCT-2001; 2001US-0330142P.
 PR 18-OCT-2001; 2001US-0330309P.
 PR 22-OCT-2001; 2001US-0341058P.
 PR 24-OCT-2001; 2001US-0339266P.
 PR 24-OCT-2001; 2001US-0343629P.
 PR 29-OCT-2001; 2001US-0349575P.
 PR 01-NOV-2001; 2001US-0346357P.
 PR 17-APR-2002; 2002US-0373260P.
 PR 19-APR-2002; 2002US-0373815P.
 PR 19-APR-2002; 2002US-0373817P.
 PR 19-APR-2002; 2002US-0373826P.
 PR 19-APR-2002; 2002US-0373884P.
 PR 22-APR-2002; 2002US-0374977P.
 PR 16-MAY-2002; 2002US-0381038P.
 PR 16-MAY-2002; 2002US-0381038P.
 PR 17-MAY-2002; 2002US-0381542P.
 PR 28-MAY-2002; 2002US-0383556P.
 PR 29-MAY-2002; 2002US-0383831P.
 PR 25-JUN-2002; 2002US-0391335P.
 PR 01-OCT-2002; 2002US-00262511.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
 PI Ott T, Gorman L, Zehusen BD, Anderson DW, Zhong M, Catterton B;
 PI Ji W, Miller CE, Raestelli L, Stone DJ, Pena CEA, Shenoy SG;
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, DiIppio VA;
 PI Eisen AV, Gangoli EA, Rieger DK, Spaderna SK;
 XX
 XX WPI; 2003-381626/36.
 DR N-PSDB; ADA05823.
 DR
 XX New NOVX polypeptides and nucleic acids, useful for diagnosing,
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or

PT pharmacogenomics.

XX Claim 1, Page 268, 586pp, English.

CC The present invention describes NOVX proteins, where X can be 1 to 55
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
 CC described above and a carrier; (2) a kit comprising, in one or more
 CC containers, the composition described above; (3) an isolated nucleic acid
 CC molecule which encodes a NOVX protein of the invention; (4) a vector
 CC comprising the nucleic acid molecule described above; (5) a cell
 CC comprising the above vector; (6) an antibody that immunospecifically
 CC binds to the polypeptide described above; (7) methods for determining the
 CC presence or amount of the above polypeptide or nucleic acid molecule in a
 CC sample; (8) methods for determining the presence of or predisposition to
 CC a disease associated with altered levels of expression of the above
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
 CC method of identifying an agent that binds to the polypeptide described
 CC above; (10) a method for identifying a potential therapeutic agent for
 CC use in treating a pathology that is related to an aberrant expression or
 CC aberrant physiological interactions of the polypeptide; (11) a method of
 CC screening for a modulator of activity or of latency or predisposition to
 CC a pathology associated with the polypeptide; (12) a method for modulating
 CC the activity of the polypeptide described above; (13) methods of treating
 CC or preventing a pathology associated with the above polypeptide in a
 CC mammal; and (14) a method for producing the above polypeptide. NOVX
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,
 CC immunomodulator, cytosstatic, neurotropic, neuroprotective, antiparkinsonian
 CC and antihypertensive activities, and can be used in gene therapy. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease. The polypeptide or the nucleic
 CC acid molecule may be used to diagnose, treat or prevent metabolic
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
 CC disease, immune disorders, haematopoietic disorders and various
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. The present sequence represents a human NOVX from the
 CC present invention.

XX Sequence 218 AA;

Query Match 71.8%; Score 1167; DB 6; Length 218;

Best Local Similarity 100.0%; Pred. No. 6,1e-113;

Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 MMSGGLAEIDSAECVFRMNAQPTVGFADVGQSTLRVSHTSVPLLRNYSHPQAR 144
 DB 1 MMSGGLAEIDSAECVFRMNAQPTVGFADVGQSTLRVSHTSVPLLRNYSHPQAR 60
 QY 145 DTLVWVGQGRHMDRVCGRTYRTLLQITRMYPGLQYTFPERMAYCDQIFQDETGR 204
 DB 61 DTLVWVGQGRHMDRVCGRTYRTLLQITRMYPGLQYTFPERMAYCDQIFQDETGR 120
 QY 205 ROSSGFSTGMPFTILALCEELVYVGMVSDSYCREKSHPSVYHYFEGKRLDECOMYL 264
 DB 121 ROSSGFSTGMPFTILALCEELVYVGMVSDSYCREKSHPSVYHYFEGKRLDECOMYL 180
 QY 265 AHEQAPRSARHFTTEKAVFSRWAKKRPVFAHPSWRTE 302
 DB 181 AHEQAPRSARHFTTEKAVFSRWAKKRPVFAHPSWRTE 218

RESULT 10

ID ADN62987 standard; protein; 218 AA.

XX ADN62987;

XX 01-JUL-2004 (first entry)

XX Human NOV44a.

XX human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;

KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;
 KW Alzheimer's disease; Parkinson's disease; immune disorder;
 KW hemetopietic disorder; dyslipidaemia; metabolic syndrome X;
 KW wasting disorder.

XX Homo sapiens.

PN US2004038223-A1.

PD 26-FEB-2004.

XX 01-OCT-2002; 2002US-00262511.

XX 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339266P.

PR 24-OCT-2001; 2001US-0343629P.

PR 29-OCT-2001; 2001US-0349575P.

PR 01-NOV-2001; 2001US-0346357P.

PR 17-APR-2002; 2002US-037360P.

PR 19-APR-2002; 2002US-0373815P.

PR 19-APR-2002; 2002US-0373817P.

PR 19-APR-2002; 2002US-0373826P.

PR 19-APR-2002; 2002US-0373844P.

PR 22-APR-2002; 2002US-0374977P.

PR 16-MAY-2002; 2002US-0381037P.

PR 16-MAY-2002; 2002US-0381038P.

PR 16-MAY-2002; 2002US-0381042P.

PR 17-MAY-2002; 2002US-0381642P.

PR 28-MAY-2002; 2002US-0383656P.

PR 29-MAY-2002; 2002US-0383831P.

PR 25-JUN-2002; 2002US-0391335P.

XX (SMIT/) SMITHSON G.

PA (MILL/) MILLET I.

PA (PEYM/) PEYMAN J A.

PA (KEKU/) KEKUDA R.

PA (JUJU/) JU J.

PA (LIIL/) LI I.

PA (GUOX/) GUO X.

PA (PATY/) PATTURAJAN M.

PA (SEPT/) SPYTEK K A.

PA (EDIN/) EDINGER S R.

PA (ELLE/) ELLERMAN K.

PA (MALY/) MALYANKAR U M.

PA (ORTT/) ORT T.

PA (GORM/) GORMAN L.

PA (ZERR/) ZERRHUSEN B D.

PA (ANDR/) ANDERSON D W.

PA (ZHON/) ZHONG M.

PA (CATT/) CATTERTON E.

PA (JTW/) JI W.

PA (MILL/) MILLER C E.

PA (RAST/) RASTELT L.

PA (STON/) STONE D J.

PA (PENA/) PENNA C E A.

PA (SHEN/) SHENOY S G.

PA (SHIM/) SHIMKETS R A.

PA (ROTH/) ROTHENBERG M E.

PA (LEAC/) LEACH M D.

PA (AGEE/) AGEER M L.

PA (BERG/) BERGHS C.

PA (DIP/) DIPIPPO V A.

PA (EISE/ EISEN A.
PA (GANG/) GANGOLLI E A.
PA (RIEG/) RIEGER D K.
PA (SPAD/) SPADERNA S K.
XX
PI Smattheon G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Pattarajan M, Spytek KA, Edinger SR, Eilerman K, MalYankar UM;
PI Orr T, Gorman L, Zechusen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEa, Shenoy SG;
PI Shinkets RA, Rothenberg ME, leach MD, Agee ML, Berghs C, Dipalpo VA,
PI Eissen A, Gangolli EA, Rieger DK, Spaderna SK;
XX
DR WPI; 2004-213931/20.
DR N-PSDB; ADN62986.
XX
XX Isolated NOXV polypeptides and nucleic acids, useful for preventing,
PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.
XX Claim 1; SEQ ID NO 184; 395pp; English.

DB 12 AVSFALCIL-----LAMRLANDVTPFLNLNCFGPKTKWIPLSYTLRQPLQTHYGV 64
 QY 60 SSVPDGRPLVPRPCSCAVVSSGQMLGSLGAEIDSACVFPFMMNQAPTVGFADYQGRS 119
 DB 65 INRTQEPFLQLNCNCAVAVNSGQWGVQKVGEEIDRASCIWMNNAPTFGFEEDYGM 123
 QY 120 TLAVHTSVPLLLRMYSHYFOKARDPLVWVGQGRHMDVLDGRYRTLLQTLTRYPG 179
 DB 124 MVRVSHSVPLLLKMPDFFKASSTTVYVWGFPRMNRKQNGIYNNMLKTVADVPDA 183
 QY 180 QVTFTEFMMAYCDQIFODETGKRRQSGSFLSTGWFMTLALCEIIVVGVNSDYSY 239
 DB 184 QIVVTTEGRTYCDGVFKDGTGKORVSSYLSGTGWFITLMDACYSIHVGMINETY 243
 QY 240 REKSHPSVPYHYEKGRLDECOMYLAEQAPRSARHFTTKAVPSRMKKRPVFAHPSW 299
 DB 244 TTGGYKRVPHYHYEKGK-DECNENYVLHEHAPYGQHRFITEKVFAMKAKGIRVFTHPM 302

RESULT 12
 AAE21578
 ID AAE21578 standard; protein; 305 AA.
 AC AAE21578;
 XX
 DT 16-JUL-2002 (first entry)
 DE Human gene 1 encoded secreted protein HE9PR39, SEQ ID NO:50.
 XX
 KM Human; secreted protein; proliferative disorder; cancer; tumour; AIDS;
 KM foetal abnormality; developmental; haematopoietic disorder; kidney; skin;
 KM immune system; cardiovascular; angiogenic; gastrointestinal; endocrine;
 KM pregnancy-related disorder; cognitive; neurological; autoimmune disease;
 KM inflammation; allergy; rheumatoid arthritis; Alzheimer's disease; asthma;
 KM Parkinson's disease; schizophrenia; psoriasis; sepsis; atherosclerosis;
 KM diabetes; infection; wound healing; vulnery; chemotaxis; food additive;
 KM gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide 1. .30
 FT Protein /label= Signal_peptide
 FT /label= Mature_secreted_protein
 FT
 XX MO200222654-A1.
 XX
 PD 21-MAR-2002.
 XX
 PF 17-JAN-2001; 2001WO-US001385.
 XX
 PR 12-SEP-2000; 2000US-0231969P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR;
 PI Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;
 PI Fiscella M, Ni J;
 XX
 DR WPI; 2002-3155684/35.
 DR N-PSDB; AAD34106.
 XX
 XX Seventeen nucleic acid molecules encoding human secreted proteins, useful
 PT for treating and preventing cancer, immune disorders (e.g. Addison's
 PT disease, and allergies), and cardiovascular disorders (e.g. myocardial
 PT ischaemia).
 XX
 PS Claim 11; Page 437-438; 483pp; English.
 XX
 CC AAD34106-AA034144 represent cDNAs corresponding to 17 human secreted
 CC protein genes, and AAE21578-AAE21616 represent the proteins they encode.
 CC AAE21617-AAE21627 represent human secreted protein fragments. The genes

CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 17 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of cancer,
 CC proliferative disorders, tumours, foetal and developmental abnormalities,
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 CC diseases (rheumatoid arthritis), inflammation, allergies, neurological
 CC disorders (Alzheimer's disease, Parkinson's disease), cognitive disorder,
 CC schizophrenia, asthma, skin disorders (psoriasis), sepsis, diabetes,
 CC atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney
 CC disorders, gastrointestinal disorders, pregnancy-related disorders,
 CC endocrine disorders, and infections. The proteins can also be used to aid
 CC wound healing and epithelial cell proliferation, to prevent skin aging
 CC due to sunburn, to maintain organs before transplantation, for supporting
 CC cell culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties. The
 CC present sequence represents a human secreted protein of the invention
 CC
 XX
 SQ Sequence 305 AA;
 XX
 Query Match 44.1%; Score 716.5; DB 5; Length 305;
 Best Local Similarity 52.7%; Pred. No. 16-65;
 Matches 129; Conservative 48; Mismatches 65; Indels 3; Gaps 3;
 QY 55 HPSGVSSVPDGRPLVPRPCSCAVVSSGQMLGSLGAEIDSACVFPFMMNQAPTVGFAD 114
 DB 61 HY-GYINVTQKQEPLOLD-CDLCAIVNSGQWGVQKVGNEIDRSSCIWMNNAPTKGYEED 118
 QY 115 VQGRSLRVYSHRSVPLLRMYSHYFOKARDPLVWVGQGRHMDVLDGRYRTLLQTLR 174
 DB 119 VGMVTLRVYSHRSVPLLRMYSHYFOKARDPLVWVGQGRHMDVLDGRYRTLLQTLR 178
 QY 175 MYPGLQVTFTEFMMAYCDQIFODETGKRRQSGSFLSTGWFMTLALCEIIVVGVNS 234
 DB 179 IYFNAQIYVTEKGRMYSYCDGVFKETGKORVSSYLSGTGWFITLMDACYSIHVGM 238
 QY 235 SDSYCREKSHPSVPYHYEKGRLDECOMYLAEQAPRSARHFTTKAVPSRMKKRPV 294
 DB 239 NDTYCKTEGKRVPHYHYEKGK-DECDYFVLHEHAPYGQHRFITEKVFAMKAKGIRIIF 297
 QY 295 AHPSW 299
 DB 298 THPNW 302

RESULT 13
 AAE21595
 ID AAE21595 standard; protein; 305 AA.
 AC AAE21595;
 XX
 DT 16-JUL-2002 (first entry)
 DE Human gene 1 encoded secreted protein HE9PR39, SEQ ID NO:67.
 XX
 KM Human; secreted protein; proliferative disorder; cancer; tumour; AIDS;
 KM foetal abnormality; developmental; haematopoietic disorder; kidney; skin;
 KM immune system; cardiovascular; angiogenic; gastrointestinal; endocrine;
 KM pregnancy-related disorder; cognitive; neurological; autoimmune disease;
 KM inflammation; allergy; rheumatoid arthritis; Alzheimer's disease; asthma;
 KM Parkinson's disease; schizophrenia; psoriasis; sepsis; atherosclerosis;
 KM diabetes; infection; wound healing; vulnery; chemotaxis; food additive;
 KM gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide 1. .30
 FT Protein /label= Signal_peptide
 FT /label= Mature_secreted_protein
 FT

FT Protein 31..305
 FT /note="Human mature secreted protein"
 XX
 XX W0200222654-A1.
 XX
 XX 21-MAR-2002.
 XX
 XX 17-JAN-2001; 2001WO-US001385.
 XX
 XX 12-SEP-2000; 2000US-0231969P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Komatsu GA, Baker KP, Birse CE, Soppet DR;
 PI Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;
 PI Flscella M, Ni U;
 XX
 XX WPI; 2002-315684/35.
 DR N-PSDB; AAD34123.
 XX
 PT Seventeen nucleic acid molecules encoding human secreted proteins, useful
 PT for treating and preventing cancer, immune disorders (e.g. Addison's
 PT disease, and allergies), and cardiovascular disorders (e.g. myocardial
 PT ischemias).
 XX
 PS Claim 11; Page 450-451; 483pp; English.
 XX
 CC AAD34106-AAD34114 represent cDNAs corresponding to 17 human secreted
 CC protein genes, and AAE21578-AAE21616 represent the proteins they encode.
 CC AAE21617-AAE21627 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 17 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of cancer,
 CC proliferative disorders, tumours, foetal and developmental abnormalities,
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 CC diseases (rheumatoid arthritis), inflammation, allergies, neurological
 CC disorders (Alzheimer's disease, Parkinson's disease), cognitive disorder,
 CC schizophrenia, asthma, skin disorders (psoriasis), sepsis, diabetes,
 CC atherosclerosis, cardiovascular disorders, angiodent disorders, kidney
 CC disorders, gastrointestinal disorders, pregnancy-related disorders,
 CC endocrine disorders, and infections. The proteins can also be used to aid
 CC wound healing and epithelial cell proliferation, to prevent skin aging
 CC due to sunburn, to maintain organs before transplantation, for supporting
 CC cell culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties. The
 CC present sequence represents a human secreted protein of the invention
 XX
 XX Sequence 305 AA;
 SQ
 Query Match 44.1%; Score 716.5; DB 5; Length 305;
 Best Local Similarity 52.7%; Pred. No. 1e-65;
 Matches 129; Conservative 48; Mismatches 65; Indels 3; Gaps 3;
 QY 55 HFGSSVSDGKRLVPEPCSCAVSSSGMGLGALIDAEVCYFRMNAQPTVGFED 114
 DB 61 HY-GYINVTQEPLOLD-CDLCAIVSNGQWGVQKNEIDRSSCIWRMNNAPTKGYED 118
 QY 115 VGRSTLRVSVSHSVPPLLNRYSHYFOKARDTLVWVGQGRHMDRVLGSGRTYTLQLTR 174
 DB 119 VGRMTLRVSVSHSVPPLLNRYSHYFOKARDTLVWVGQGRHMDRVLGSGRTYTLQLTR 178
 QY 175 MYEGQLVYTFERMAYCDQIFQDETCKNRSSGFLSTGFTMTIALAECEIIVYGV 234
 DB 179 IYPAQIYVTTKMSYCDGVFKKETGKDRVSGSYSTGFTMTIALMDQYGIHYGMI 238
 QY 235 SDSTCKREKSHPSVDYHYFEKGRLEDCOMYLAHQAPSAHFRTTEKAVFSRMAKRPITV 294
 DB 239 NDTYCKTEGYRKVYHYFEQGR-DECDYFLHEHAHYGGRFTEKKVPAKMAKRIIF 297

QY 295 AHPSW 299
 DB 298 TTHPW 302
 RESULT 14
 ID ABG65165 standard; protein; 305 AA.
 XX
 XX ABG65165;
 AC
 XX
 XX 27-AUG-2002 (first entry)
 DE
 XX
 XX Human albumin fusion protein #1840.
 XX
 XX Albumin fusion protein; therapeutic protein X; human albumin; HA;
 KW human serum albumin; HSA; cancer; reproductive disorder;
 KW digestive disorder; immune disorder; endocrine disorder;
 KW haematopoietic disorder; neural disorder; connective disorder;
 KW cytostatic; antifertility; antiinflammatory; anticancer;
 KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
 KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
 KW osteopathic; antiarthritic.
 XX
 XX Homo sapiens.
 OS
 XX Synthetic.
 XX
 XX W0200177137-A1.
 XX
 XX 18-OCT-2001.
 PD
 XX
 XX 12-APR-2001; 2001WO-US011988.
 PF
 XX
 XX 12-APR-2000; 2000US-0229358P.
 PR
 XX 25-APR-2000; 2000US-019384P.
 PR 21-DEC-2000; 2000US-0256931P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Haseltine WA;
 PI WPI; 2002-010886/01.
 DR
 XX
 XX New fusion protein for treating disease e.g. diabetes comprises an
 PT albumin fused to a therapeutic protein.
 PT
 XX
 XX Claim 1; Page 1796-1797; 2102pp; English.
 PS
 XX The present invention relates to albumin fusion proteins comprising a
 CC therapeutic protein X and human albumin (HA, also known as human serum
 CC albumin, HSA). The proteins are useful for treating a disease or disorder
 CC that may be modulated by therapeutic protein X. The albumin extends the
 CC shelf-life of protein X, and may increase its biological in vitro/in vivo
 CC activity. The protein is useful for treating and diagnosing disorders
 CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
 CC disease, ulcerative colitis), immune disorders (e.g. acquired
 CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
 CC haematopoietic disorders, neural disorders (e.g. Alzheimer's,
 CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,
 CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).
 CC ABG65326-ABG65518 represent albumin fusion proteins of the invention
 XX
 XX Sequence 305 AA;
 SQ
 Query Match 44.1%; Score 716.5; DB 5; Length 305;
 Best Local Similarity 52.7%; Pred. No. 1e-65;
 Matches 129; Conservative 48; Mismatches 65; Indels 3; Gaps 3;
 QY 55 HFGSSVSDGKRLVPEPCSCAVSSSGMGLGALIDAEVCYFRMNAQPTVGFED 114
 DB 61 HY-GYINVTQEPLOLD-CDLCAIVSNGQWGVQKNEIDRSSCIWRMNNAPTKGYED 118

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2004, 18:04:28 ; Search time 192 Seconds

(Without alignments)
905.016 Million cell updates/sec

Title: US-10-019-114A-7

Perfect score: 1626

Sequence: 1 MKAPGRVLIIICSVFSAV.....FSRWAKKRPIVFAHPMSWRT 302

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1626	100.0	302	1	SI7D HUMAN
2	1597	98.2	302	2	Q6ZXX1 h alpha-n-a
3	1597	98.2	302	2	Q6ZXX1 pan troglod
4	1479	91.0	302	2	Q704X3 sus scrofa
5	1479	91.0	302	2	Q704X3 bos taurus
6	1476	90.8	323	2	Q704X4
7	1476	90.8	323	2	Q704X4
8	1452	89.3	302	2	Q6C3J2 m mus muscu
9	1452	89.3	317	2	Q6PHP9
10	1452	89.3	317	2	AAH56451
11	1439	88.5	360	1	SI7D MOUSE
12	1436	88.3	302	2	Q6ZXX0
13	1436	88.3	302	2	Q6ZXX0
14	1047.5	64.4	289	2	Q704X2
15	1047.5	64.4	289	2	Q704X2
16	740	45.5	306	2	Q704S4
17	740	45.5	306	2	Q704S4
18	740	45.5	306	2	Q704S4
19	733.5	45.1	305	1	SI7C MOUSE
20	733.5	45.1	305	1	SI7C MOUSE
21	728.5	44.8	305	2	Q6IN13
22	728.5	44.8	305	2	AAH75501
23	726.5	44.7	305	1	SI7C RAT
24	726.5	44.7	299	2	Q6ZXX0
25	725.5	44.6	299	2	Q6ZXX0
26	724	44.5	292	2	Q6ZXX2
27	724	44.5	292	2	Q6ZXX2
28	718.5	44.2	300	2	Q6ZXX2
29	718.5	44.2	300	2	Q6ZXX2
30	717	44.1	234	2	Q9W6U6
31	716.5	44.1	305	2	Q6PC60

32	716.5	44.1	305	2	AAH59363	AAH59363 homo sapi
33	715.5	44.0	305	2	Q6ZXX1	Q6ZXX1 pan troglod
34	715.5	44.0	305	2	Q6ZXX1	Q6ZXX1 pan troglod
35	714.5	43.9	305	1	SI7C HUMAN	Q6ZXX1 homo sapien
36	682	41.9	138	2	Q6ZXX2	Q6ZXX2 mus muscu
37	652	40.1	311	2	Q6ZXX7	Q6ZXX7 brachydanio
38	652	40.1	311	2	Q6ZXX7	Q6ZXX7 brachydanio
39	639.5	39.3	335	2	Q6ZXX6	Q6ZXX6 pan troglod
40	639.5	39.3	335	2	Q6ZXX6	Q6ZXX6 pan troglod
41	639.5	39.3	336	1	SI7E HUMAN	Q6ZXX7 homo sapien
42	639.5	39.3	336	2	CAD45372	CAD45372 homo sapi
43	630	38.7	335	2	Q6CAM7	Q6CAM7 m mus muscu
44	629	38.7	335	2	AAH55737	AAH55737 mus muscu
45	629	38.7	335	2	BAC31331	BAC31331 m 7 days

ALIGNMENTS

RESULT 1
SI7D HUMAN STANDARD; PRT; 302 AA.
ID Q9H4FL; Q9NMU6; Q9UKU1; Q9ULB9; Q9Y3G3; Q9Y3G4;
AC Q9H4FL; Q9NMU6; Q9UKU1; Q9ULB9; Q9Y3G3; Q9Y3G4;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 05-JUN-2004 (Rel. 44, Last annotation update)
DE Alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3-N-acetyl-
DE galactosaminide alpha-2,6-sialyltransferase (EC 2.4.99.7) (NeuAc-
DE alpha-2,3-Gal-beta-1,3-GalNAc-alpha-2,6-sialyltransferase) (St6GalNAc
DE IV) (Sialyltransferase 7D) (Sialyltransferase 3C).
OS Homo sapiens (Human).
GN Name=SIAT7D; Synonyms=SIAT3C;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20517254; PubMed=11062056;
RA Harduin-lepers A., Stokes D.C., Steelant W.F.A., Samyn-Petit B.,
RA Krzewinski-Rechni M.A., Vallejo-Ruiz V., Zanetta J.P., Auge C.,
RA Delannoy P.;
RT "Cloning, expression and gene organization of a human Neu5Ac alpha-2-
RT 3Gal beta 1-3GalNAc alpha 2,6-sialyltransferase: hST6GalNAcIV";
RL Biochem. J. 352:337-348(2000). November issue
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal liver;
RA Kim K.-W., Kim K.-S., Do S.-I., Kim C.-H., Lee Y.-C.;
RT "Molecular cloning of Neu5Ac alpha 2,6-sialyltransferase (ST6GalNAc alpha 2,6-
RT sialyltransferase) cDNA from human fetal liver";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=lung;
RA Yoshida A.;
RT "N-acetyl-galactosaminide alpha2,6-sialyltransferase";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahori K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuna M.,
RA Shigahori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuma N., Sato K., Tanikawa M.,
RA Yamazaki M., Nishimura K., Iwibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
RA Ishigori S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Houchi T.,
RA Kusano Y., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiyama S., Komai F., Hara R., Takeuchi K., Arita M.,

RA Imose N., Musashino K., Yuuki H., Oshima A., Saeki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Matanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Matanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujizawa T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohnori Y.,
 RA Kawabata A., Hiki T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Ohtsuki R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senda T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Matanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Maeno Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isozaki T., Sugano S.,
 RT Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs. Nat. Genet. 36:40-45(2004).
 RL
 RN
 RP
 RC
 RX MEDLINE=2368257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusik A., Farmer A.A., Rubin G.M., Hong L.,
 RA Saplinton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Umeda T.B., Toshikiy S., Cantiani P., Prange C.,
 RA Rasmussen S., Lomquist N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S., Loguanello N.A., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.D., Hultk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fehey J., Hailton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Rodriguez R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Brakley A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalek U., Small D.E.,
 RA Schmeich A., Schein J.E., Jones S.U.M., Marra M.A.,
 RT Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences. Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RL
 RN
 RP
 RX MEDLINE=9929277; PubMed=10369878;
 RA Gilley J., Fried M.,
 RT Extensive gene order differences within regions of conserved synteny
 RT between the Fugu and human genomes: implications for chromosomal
 RT evolution and the cloning of disease genes. Hum. Mol. Genet. 8:1333-1320(1999).
 RL
 CC -1- FUNCTION: Involved in the biosynthesis of ganglioside GD1a from
 CC GM1b. Transfers CMP-NeuAc with an alpha-2,6-linkage to GalNAc
 CC residue on NeuAc-alpha-2,3-Gal-beta-1,3-GalNAc of glycoproteins
 CC and glycolipids. Prefers glycoproteins to glycolipids (By
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: CMP-N-acetylneuraminic acid + N-acetyl-alpha-
 CC neuraminyl-(2->3)-beta-D-galactosyl-(1->3)-N-acetyl-D-
 CC galactosaminyl-R = CMP + N-acetyl-alpha-neuraminyl-(2->3)-beta-D-
 CC galactosyl-(1->3)-[N-acetyl-alpha-neuraminyl-(2->6)]-N-acetyl-D-
 CC galactosaminyl-R.
 CC -1- PATHWAY: Glycosylation.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (By
 CC similarity).
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase family 29.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

CC
 DR EMBL: A0271734; CAC07404.1; -
 DR EMBL: AF127142; AAF00102.1; -
 DR EMBL: AB035172; BAA87034.1; -
 DR EMBL: AK006000; BAA91281.1; -
 DR EMBL: BC037005; AAB36705.1; -
 DR EMBL: Y17460; CAB44353.1; -
 DR EMBL: Y17461; CAB44354.1; -
 DR GenBank: H09C117846; STAT7D.
 DR
 DR InterPro: IPR001675; Glyco trans_29.
 DR Pfam: PF00777; Glyco transf_29; 1.
 KW Glycoprotein; Glycosyltransferase; Golgi stack; Signal-anchor;
 KW Transferase; Transmembrane.
 FT DOMAIN 1 6
 FT TRANSMEM 7 27
 FT
 FT DOMAIN 28 302
 FT DISULFID 76 225
 FT CARBOHYD 135 135
 FT CONFLICT 119 119
 FT CONFLICT 119 120
 FT CONFLICT 140 140
 SQ SEQUENCE 302 AA; 34200 MW; 08A4CD749A6D783 CRC64;
 Query Match 100.0%; Score 1626; DB 1; Length 302;
 Best Local Similarity 100.0%; Pred. No. 1.3e-139; Indels 0; Gaps 0;
 Matches 302; Conservative 0; Mismatches 0;
 QY 1 MKAPGLVLLILCSVFSNAVYLLCCWAGLPLCLATCLDHPPTGSRPTVPGELHFGYS 66
 DB 1 MKAPGLVLLILCSVFSNAVYLLCCWAGLPLCLATCLDHPPTGSRPTVPGELHFGYS 60
 QY 61 SVDPGRPLVREPRSCAVVSSSGMGSIGAIIDBAECYFRNNQAPTYGFADVQRST 120
 DB 61 SVDPGRPLVREPRSCAVVSSSGMGSIGAIIDBAECYFRNNQAPTYGFADVQRST 120
 QY 121 LRVVSHTSVPLLRNYSHPYQKARDTLVYWGGRMDRVLGSRVYRTLLQLTRMPSGLQ 180
 DB 121 LRVVSHTSVPLLRNYSHPYQKARDTLVYWGGRMDRVLGSRVYRTLLQLTRMPSGLQ 180
 QY 181 VYTFERMAVCCQIFQDETGNKRRSGSFLSTGWTMLALCEIIVYGVSDSYCR 240
 DB 181 VYTFERMAVCCQIFQDETGNKRRSGSFLSTGWTMLALCEIIVYGVSDSYCR 240
 QY 241 EKSHPSVPHYEFKGLDEQWTLAEQAPRSNRHRTTEKAVSSRAKRPVFAHPSWR 300
 DB 241 EKSHPSVPHYEFKGLDEQWTLAEQAPRSNRHRTTEKAVSSRAKRPVFAHPSWR 300
 QY 301 TE 302
 DB 301 TE 302
 RESULT 2
 06ZXZ1
 ID 06ZXZ1 PRELIMINARY; PRT; 302 AA.
 AC 06ZXZ1;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Alpha-2,6-sialyltransferase ST6GALNAc IV (Fragment).
 GN Name=stat7d;
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
 OX NCBI_TaxID=9598;
 RN
 RP SEQUENCE FROM N.A.
 RA Hardin-Lepers A., Martinez-Duncker I., Mollicone R., Delannoy P.,
 RA Orici R.,
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: A046870; CAG26699.1; -

DR GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
 DR InterPro: IPR001675; Glyco_trans_29.
 DR Pfam: PF00777; Glyco_trans_29; 1.
 KW Glycosyltransferase; transferase.
 FT NON TER 1
 FT NON TER 302
 SQ SEQUENCE 302 AA; 34138 MW; AB7C984ABE53B553 CRC64;

Query Match 98.2%; Score 1597; DB 2; Length 302;
 Best Local Similarity 98.0%; Pred. No. 5.5e-137;
 Matches 296; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKAPGRVLVILICSVPSAVYIILCCWAGPLCLATCLDHHPTGSRPTVGPPLHFGYS 60
 DB 1 MKASGRVLVILICSVPSAVYIILCCWASLPLCLATCLDHHPTGSRPTVGPPLHFGYS 60
 QY 61 SVDPGKPLVREPCSCAVVSSSGOMLGGAGLIDSAECVFRMNOAPTVGFADVGQRST 120
 DB 61 SVDPGKPLVREPCSCAVVSSSGOMLGGAGLIDSAECVFRMNOAPTVGFADVGQRST 120
 QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLAGRTYRTLLQLTRMYPGLQ 180
 DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLAGRTYRTLLQLTRMYPGLQ 180
 QY 181 VYFTEFRMAVCCQIFODETGKRRSGSFLSTGTFMTIALBELCEIIVYGVMSDSYCR 240
 DB 181 VYFTEFRMAVCCQIFODETGKRRSGSFLSTGTFMTIALBELCEIIVYGVMSDSYCR 240
 QY 241 EKSHPSVPHYFEKGRGLDECOMYLAEQAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300
 DB 241 EKSHPSVPHYFEKGRGLDECOMYLAEQAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300
 QY 301 TE 302
 DB 301 TE 302

QY 301 TE 302
 DB 301 TE 302

RESULT 3
 CAG26699 PRELIMINARY; PRT; 302 AA.
 AC CAG26699;
 DT 23-APR-2004 (TrEMBLrel. 27, Created)
 DT 23-APR-2004 (TrEMBLrel. 27, Last sequence update)
 DE Alpha-2,6-sialyltransferase ST6GalNAc IV (Fragment).
 GN SIATD.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harduin-Leperc A., Martinez-Duncker I., Mollicone R., Delannoy P.,
 RA Oriol R.;
 RT "Phylogeny of sialyltransferases";
 RL Submitted (APR-2004) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AJ646870; CAG26699.1; -;
 KW Glycosyltransferase; transferase.
 FT NON TER 1
 FT NON TER 302
 SQ SEQUENCE 302 AA; 34138 MW; AB7C984ABE53B553 CRC64;

Query Match 98.2%; Score 1597; DB 2; Length 302;
 Best Local Similarity 98.0%; Pred. No. 5.5e-137;
 Matches 296; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKAPGRVLVILICSVPSAVYIILCCWAGPLCLATCLDHHPTGSRPTVGPPLHFGYS 60
 DB 1 MKASGRVLVILICSVPSAVYIILCCWASLPLCLATCLDHHPTGSRPTVGPPLHFGYS 60
 QY 61 SVDPGKPLVREPCSCAVVSSSGOMLGGAGLIDSAECVFRMNOAPTVGFADVGQRST 120
 DB 61 SVDPGKPLVREPCSCAVVSSSGOMLGGAGLIDSAECVFRMNOAPTVGFADVGQRST 120

QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLAGRTYRTLLQLTRMYPGLQ 180
 DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLAGRTYRTLLQLTRMYPGLQ 180
 QY 181 VYFTEFRMAVCCQIFODETGKRRSGSFLSTGTFMTIALBELCEIIVYGVMSDSYCR 240
 DB 181 VYFTEFRMAVCCQIFODETGKRRSGSFLSTGTFMTIALBELCEIIVYGVMSDSYCR 240
 QY 241 EKSHPSVPHYFEKGRGLDECOMYLAEQAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300
 DB 241 EKSHPSVPHYFEKGRGLDECOMYLAEQAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300
 QY 301 TE 302
 DB 301 TE 302

QY 301 TE 302
 DB 301 TE 302

RESULT 4
 ID Q704X3 PRELIMINARY; PRT; 302 AA.
 AC Q704X3;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE Beta-galactosamide alpha-2,6-sialyltransferase (EC 2.4.99.7).
 GN Name=siat7b; Synonyms=ST6GALNAc-IV;
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harduin-Leperc A., Martinez-Duncker I., Mollicone R., Delannoy P.,
 RA Oriol R.;
 RL Submitted (JAN-2004) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AJ620652; CAF05851.1; -;
 DR GO:0047290; F:(alpha-N-acetylneuraminyl)-2,3-beta-galactosyl. . .; IEA.
 DR GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
 DR InterPro: IPR001675; Glyco_trans_29.
 DR Pfam: PF00777; Glyco_transf_29; 1.
 KW Glycosyltransferase; transferase.
 FT NON TER 1
 FT NON TER 302
 SQ SEQUENCE 302 AA; 34226 MW; 6C48E3977DA96733 CRC64;

Query Match 91.0%; Score 1479; DB 2; Length 302;
 Best Local Similarity 90.1%; Pred. No. 3.1e-126;
 Matches 272; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 1 MKAPGRVLVILICSVPSAVYIILCCWAGPLCLATCLDHHPTGSRPTVGPPLHFGYS 60
 DB 1 MKAPGRVLVILICSVPSAVYIILCCWAGPLCLATCLDHHPTGSRPTVGPPLHFGYS 60
 QY 61 SVDPGKPLVREPCSCAVVSSSGOMLGGAGLIDSAECVFRMNOAPTVGFADVGQRST 120
 DB 61 SVDPGKPLVREPCSCAVVSSSGOMLGGAGLIDSAECVFRMNOAPTVGFADVGQRST 120
 QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLAGRTYRTLLQLTRMYPGLQ 180
 DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRVLAGRTYRTLLQLTRMYPGLQ 180
 QY 181 VYFTEFRMAVCCQIFODETGKRRSGSFLSTGTFMTIALBELCEIIVYGVMSDSYCR 240
 DB 181 VYFTEFRMAVCCQIFODETGKRRSGSFLSTGTFMTIALBELCEIIVYGVMSDSYCR 240
 QY 241 EKSHPSVPHYFEKGRGLDECOMYLAEQAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300
 DB 241 EKSHPSVPHYFEKGRGLDECOMYLAEQAPRSARHRTTEKAVFSRMAKKRPIVFAHPSWR 300
 QY 301 TE 302
 DB 301 TE 302

```

RESULT 5
ID CAF05851 PRELIMINARY; PRT; 302 AA.
AC CAF05851;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DE Beta-galactosamide alpha-2,6-sialyltransferase (EC 2.4.99.7).
GN ST6GALNAC-IV OR SIAT7B.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Hardin-Leipers A., Martinez-Duncker I., Mollicone R., Delanny P.,
RA Oriol R.;
RT "Phylogeny of sialyltransferases."
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ620652; CAF05851.1; -.
KW Glycosyltransferase; Transferase.
SQ
SEQUENCE 302 AA; 3426 MM; 6C48E3977DA96733 CRC64;

Query Match 91.0%; Score 1479; DB 2; Length 302;
Best Local Similarity 90.1%; Pred. No. 3.1e-126;
Matches 272; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIICSVFSNAVYIILCCWAGLPCLATCLDHHFPTGSRPTVPGPLHFGSGYS 60
DB 1 MKPGRLLIILICSLGSGFYLILCCWACLPFCPLATCLDPLISINSRPTVPGPLHFGSGYS 60
QY 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
DB 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
QY 121 LRVVSHTSVPLLIRNYSHYFOKARDTLVYWGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
DB 121 LRVVSHTSVPLLIRNYSHYFOKARDTLVYWGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
QY 181 VYFTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
DB 181 VYFTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
QY 241 EKSHPSVPHYFEKGRDECOMYLAEQAPRSARHFTTEKAVFSRMAKRPVFAHPSWR 300
DB 241 EKSHPSVPHYFEKGRDECOMYLAEQAPRSARHFTTEKAVFSRMAKRPVFAHPSWR 300
QY 301 TE 302
DB 301 TE 302
QY 301 TQ 302
DB 301 TQ 302

RESULT 6
ID Q704X4 PRELIMINARY; PRT; 323 AA.
AC Q704X4;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Beta-galactosamide alpha-2,6-sialyltransferase (EC 2.4.99.7).
GN Name=siat7b; Synonyms=st6galnac-iv;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Hardin-Leipers A., Martinez-Duncker I., Mollicone R., Delanny P.,
RA Oriol R.;
RT "Phylogeny of sialyltransferases."
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ620651; CAF05850.1; -.
KW Glycosyltransferase; Transferase.
SQ
SEQUENCE 323 AA; 36765 MM; 4BBF007BF6AAE481 CRC64;

Query Match 90.8%; Score 1476; DB 2; Length 323;
Best Local Similarity 91.1%; Pred. No. 6.3e-126;
Matches 275; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIICSVFSNAVYIILCCWAGLPCLATCLDHHFPTGSRPTVPGPLHFGSGYS 60
DB 1 MKAPGRVLIIICSVFSNAVYIILCCWAGLPCLATCLDHHFPTGSRPTVPGPLHFGSGYS 60
QY 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
DB 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
QY 121 LRVVSHTSVPLLIRNYSHYFOKARDTLVYWGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
DB 121 LRVVSHTSVPLLIRNYSHYFOKARDTLVYWGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
QY 181 VYFTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
DB 181 VYFTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
QY 241 EKSHPSVPHYFEKGRDECOMYLAEQAPRSARHFTTEKAVFSRMAKRPVFAHPSWR 300
DB 241 EKSHPSVPHYFEKGRDECOMYLAEQAPRSARHFTTEKAVFSRMAKRPVFAHPSWR 300
QY 301 TE 302
DB 301 TE 302
QY 301 TQ 302
DB 301 TQ 302

RESULT 7
ID CAF05850 PRELIMINARY; PRT; 323 AA.
AC CAF05850;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE Beta-galactosamide alpha-2,6-sialyltransferase (EC 2.4.99.7).
GN SIAT7B OR ST6GALNAC-IV.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Hardin-Leipers A., Martinez-Duncker I., Mollicone R., Delanny P.,
RA Oriol R.;
RT "Phylogeny of sialyltransferases."
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ620651; CAF05850.1; -.
KW Glycosyltransferase; Transferase.
SQ
SEQUENCE 323 AA; 36765 MM; 4BBF007BF6AAE481 CRC64;

Query Match 90.8%; Score 1476; DB 2; Length 323;
Best Local Similarity 91.1%; Pred. No. 6.3e-126;
Matches 275; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIICSVFSNAVYIILCCWAGLPCLATCLDHHFPTGSRPTVPGPLHFGSGYS 60
DB 1 MKAPGRVLIIICSVFSNAVYIILCCWAGLPCLATCLDHHFPTGSRPTVPGPLHFGSGYS 60
QY 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
DB 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
QY 121 LRVVSHTSVPLLIRNYSHYFOKARDTLVYWGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
DB 121 LRVVSHTSVPLLIRNYSHYFOKARDTLVYWGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
QY 181 VYFTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
DB 181 VYFTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
QY 241 EKSHPSVPHYFEKGRDECOMYLAEQAPRSARHFTTEKAVFSRMAKRPVFAHPSWR 300
DB 241 EKSHPSVPHYFEKGRDECOMYLAEQAPRSARHFTTEKAVFSRMAKRPVFAHPSWR 300
QY 301 TE 302
DB 301 TE 302
QY 301 TQ 302
DB 301 TQ 302

```

```

DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR001675; Glyco_trans_29.
DR Pfam; PF00777; Glyco_transf_29; I.
KW Glycosyltransferase; Transferase.
SQ
SEQUENCE 323 AA; 36765 MM; 4BBF007BF6AAE481 CRC64;

Query Match 90.8%; Score 1476; DB 2; Length 323;
Best Local Similarity 91.1%; Pred. No. 6.3e-126;
Matches 275; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIICSVFSNAVYIILCCWAGLPCLATCLDHHFPTGSRPTVPGPLHFGSGYS 60
DB 1 MKAPGRVLIIICSVFSNAVYIILCCWAGLPCLATCLDHHFPTGSRPTVPGPLHFGSGYS 60
QY 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
DB 61 SVDPGKPLVREPCSCAVSSSGOMLGGAGIIDAECVFRNQAFTVGFEDVQGRST 120
QY 121 LRVVSHTSVPLLIRNYSHYFOKARDTLVYWGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
DB 121 LRVVSHTSVPLLIRNYSHYFOKARDTLVYWGQGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
QY 181 VYFTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
DB 181 VYFTFERMAVYCOIFQDETKNRROSGSFLSTGFMFIATLCEIIVYGVWSDSYCR 240
QY 241 EKSHPSVPHYFEKGRDECOMYLAEQAPRSARHFTTEKAVFSRMAKRPVFAHPSWR 300
DB 241 EKSHPSVPHYFEKGRDECOMYLAEQAPRSARHFTTEKAVFSRMAKRPVFAHPSWR 300
QY 301 TE 302
DB 301 TE 302
QY 301 TQ 302
DB 301 TQ 302

```


Db 121 LRVISHTSVPLLRNNSHYFQOARDLTVVWGGCKHMDRALGGRTRYPALLQLTFRMYPGLQ 180
 QY 181 VYFTEPRMAVYCCQIFQDEGTGKRRSGSFLSTGFMFMILALCEIIVVYGWVSDSYCR 240
 Db 181 VYFTEPRMAVYCCQIFQDEGTGKRRSGSFLSTGFMFMILALCEIIVVYGWVSDSYCR 240
 QY 241 EKHPSPVHYHYFEKGRIDECCOMYLAHQAPRSARHRTTEKAVFSRMACKRPVFAHPSWR 300
 Db 241 EKHHPSPVHYHYFEKGRIDECCOMYLAHQAPRSARHRTTEKAVFSRMACKRPVFAHPSWR 300
 QY 301 TE 302
 Db 301 TQ 302

RESULT 8
 ID 08C3J2 PRELIMINARY; PRT; 302 AA.
 AC 08C3J2;
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
 DE Mus musculus 10 days lactation, adult female mammary gland cDNA, RIKEN
 DE full-length enriched library, clone:D73003H08
 DE product:beta1ytransferase 7 (alpha-N-acetylneuraminyl 2,3-
 DE beta-galactosyl-1,3)-N-acetyl galactosaminide alpha-2,6-
 DE beta1ytransferase) D, full insert sequence.
 GN Name=Stact7d;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
 RX MEDLINE=9279253; Pubmed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."; Meth. Enzymol. 303:19-44(1999).
 RL [2]
 RN [2]
 RC SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
 RX MEDLINE=21085660; Pubmed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).
 RL [3]
 RN [3]
 RC SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."; Nature 420:563-573(2002).
 RL [4]
 RN [4]
 RC SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
 RX MEDLINE=20499374; Pubmed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes."; Genome Res. 10:1617-1630(2000).
 RL [5]
 RN [5]
 RC SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
 RX MEDLINE=20530913; Pubmed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Kitenai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Makunoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujisake S., Inoue K., Togawa Y., Izawa M., Ohara E., Matsubiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,

RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer."; Genome Res. 10:1757-1771(2000).
 RL [6]
 RN [6]
 RC SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ono N., Okazaki Y.,
 RA Saito R., Saitoh K., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tegan M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK085730; BAC39523.1;
 DR MGD: MGI:1341894; Stact7d.
 DR GO: GO:0005794; C:Golgi apparatus; IEA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0008373; F:beta1ytransferase activity; IEA.
 DR GO: GO:0006486; P:protein amino acid glycosylation; IEA.
 DR InterPro: IPR001675; Glyco trans_29.
 DR Pfam: PF00777; Glyco trans_29; I.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 302 AA; 34218 MW; 3AF7FC04F39E310B CRC64;

Query Match 89.3%; Score 1452; DB 2; Length 302;

Best Local Similarity 89.1%; Pred. No. 8.9e-124; Matches 269; Conservative 9; Mismatches 24; Indels 0; Gaps 0;

QY 1 MKAPGRVLVLLICGVVFSAYVLLCCWAGLPLCLATCLDHHFPTGSRPTVGPGLHPSGYS 60
 Db 1 MKAPGRVLVLLICGVVFSAYVLLCCWAGLPLCLATCLDHHFPTGSRPTVGPGLHPSGYS 60
 QY 61 SVDPGRPLVREPCSCAVSSSGQMGIGABIDAECEYFRNQAFTVGFEDVCGRST 120
 Db 61 SVDPGRPLVREPCSCAVSSSGQMGIGABIDAECEYFRNQAFTVGFEDVCGRST 120
 QY 121 LRVISHTSVPLLRNNSHYFQOARDLTVVWGGCKHMDRALGGRTRYPALLQLTFRMYPGLQ 180
 Db 121 LRVISHTSVPLLRNNSHYFQOARDLTVVWGGCKHMDRALGGRTRYPALLQLTFRMYPGLQ 180
 QY 181 VYFTEPRMAVYCCQIFQDEGTGKRRSGSFLSTGFMFMILALCEIIVVYGWVSDSYCR 240
 Db 181 VYFTEPRMAVYCCQIFQDEGTGKRRSGSFLSTGFMFMILALCEIIVVYGWVSDSYCR 240
 QY 241 EKHPSPVHYHYFEKGRIDECCOMYLAHQAPRSARHRTTEKAVFSRMACKRPVFAHPSWR 300
 Db 241 EKHPSPVHYHYFEKGRIDECCOMYLAHQAPRSARHRTTEKAVFSRMACKRPVFAHPSWR 300
 QY 301 TE 302
 Db 301 AK 302

RESULT 9
 ID 06PH9 PRELIMINARY; PRT; 317 AA.
 AC 06PH9;
 DT 05-JUL-2004 (TEMBLrel. 27, Created)
 DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
 DE Siact7 protein (Fragment).
 GN Name=Siact7d;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywnicki M.I., Skaleka U., Smalins D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RA Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC056451; AAH56451.1;
 DR InterPro; IPR001675; Glyco trans_29.
 DR Pfam; PF00777; Glyco transf_29; 1.
 FT NON_TER
 SQ SEQUENCE 317 AA; 35740 MW; 61B8713C0234360D CRC64;

Query Match 89.3%; Score 1452; DB 2; Length 317;
 Best Local Similarity 89.1%; Pred. No. 9.5e-124;
 Matches 269; Conservative 9; Mismatches 24; Indels 0; Gaps 0;

QY 1 MKAPGRVLILICSVPSAVYIILCCWAGLPCLATCLDHHFPGSRPTVPGRLHFGYS 60
 DB 16 MKAPGRLLLTLLTCLIFSAVCVPLCCMACLPCLATCLDRHLPAPRSTVPGRLHFGYS 75
 QY 61 SVDPGKPLVREPCSCAVVSSSGOMLGSGIAGTIDSAECVFRMNOAPTYGFEDVQGRST 120
 DB 76 SVDPGKPLRELCHSCAVVSSSGOMLGSGIAGTIDGAEVLRMNOAPTYGFEDVQGRST 135
 QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRYLGRTYRTLLQLTRMYPGLQ 180
 DB 136 LRVVSHTSVPLLRNYSHYFOHARDTLVYWGGRHMDRYLGRTYRTLLQLTRMYPGLQ 195
 QY 181 VYFTFRMAVYCOIQFODETGKRRSGSFLSTGWTMTIALELCEIIVYGVNSDYSYCR 240
 DB 196 VYFTFRMAVYCOIQFODETGKRRSGSFLSTGWTMTIALELCEIIVYGVNSDYSYCS 255
 QY 241 EKSHPSVYHYFEKGRGLDECOMYLAEQAPRSARHFTTEKAVFSRWAKKRPVFAHPSWR 300
 DB 256 EKSPRSVYHYFEKGRGLDECOMYLAEQAPRSARHFTTEKAVFSRWAKKRPVFAHPSWR 315
 QY 301 TE 302
 DB 316 AK 317

RESULT 10
 AAH56451 PRELIMINARY; PRT; 317 AA.

AC AAH56451;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DE stat7d protein (Fragment).
 GN STAT7D.
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywnicki M.I., Skaleka U., Smalins D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RA Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC056451; AAH56451.1;
 DR InterPro; IPR001675; Glyco trans_29.
 DR Pfam; PF00777; Glyco transf_29; 1.
 FT NON_TER
 SQ SEQUENCE 317 AA; 35740 MW; 61B8713C0234360D CRC64;

Query Match 89.3%; Score 1452; DB 2; Length 317;
 Best Local Similarity 89.1%; Pred. No. 9.5e-124;
 Matches 269; Conservative 9; Mismatches 24; Indels 0; Gaps 0;

QY 1 MKAPGRVLILICSVPSAVYIILCCWAGLPCLATCLDHHFPGSRPTVPGRLHFGYS 60
 DB 16 MKAPGRLLLTLLTCLIFSAVCVPLCCMACLPCLATCLDRHLPAPRSTVPGRLHFGYS 75
 QY 61 SVDPGKPLVREPCSCAVVSSSGOMLGSGIAGTIDSAECVFRMNOAPTYGFEDVQGRST 120
 DB 76 SVDPGKPLRELCHSCAVVSSSGOMLGSGIAGTIDGAEVLRMNOAPTYGFEDVQGRST 135
 QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYWGGRHMDRYLGRTYRTLLQLTRMYPGLQ 180
 DB 136 LRVVSHTSVPLLRNYSHYFOHARDTLVYWGGRHMDRYLGRTYRTLLQLTRMYPGLQ 195
 QY 181 VYFTFRMAVYCOIQFODETGKRRSGSFLSTGWTMTIALELCEIIVYGVNSDYSYCR 240
 DB 196 VYFTFRMAVYCOIQFODETGKRRSGSFLSTGWTMTIALELCEIIVYGVNSDYSYCS 255
 QY 241 EKSHPSVYHYFEKGRGLDECOMYLAEQAPRSARHFTTEKAVFSRWAKKRPVFAHPSWR 300
 DB 256 EKSPRSVYHYFEKGRGLDECOMYLAEQAPRSARHFTTEKAVFSRWAKKRPVFAHPSWR 315
 QY 301 TE 302
 DB 316 AK 317

RESULT 11
 S17D MOUSE STANDARD; PRT; 360 AA.

AC Q9R2B6; O88725; Q9JHP0; Q9QUP9; Q9R2B5;
 DT 28-FEB-2003 (Ref. 41, Created)
 DT 28-FEB-2003 (Ref. 41, Last sequence update)
 DT 05-JUL-2004 (Ref. 44, Last annotation update)
 DE Alpha-N-acetyl-neuraminy-2,3-beta-galactosyl-1,3-N-acetyl-

DE galactosaminide alpha-2,6-sialyltransferase (EC 2.4.99.7) (NeuAc-
 DE alpha-2,3-Gal-beta-1,3-GalNAc-alpha-2,6-sialyltransferase) (ST6GalNAc
 DE IV) (Sialyltransferase 7D).
 GN Name=Stac7d;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A., ACTIVITY, TISSUE SPECIFICITY, ALTERNATIVE
 RP SPLICING, AND DEVELOPMENTAL STAGE.
 RC STRAIN=ICR; TISSUE=Brain;
 RX MEDLINE=9923252; PubMed=10207017;
 RA Lee Y.-C., Kaufman M., Kitzazume-Kawaguchi S., Kono M., Takashima S.,
 RA Kurosawa N., Liu H., Fletcher H., Tsuji S.;
 RT "Molecular cloning and functional expression of two members of mouse
 RT NeuAc-alpha-2,3Gal-beta-1,3GalNAc GalNAc-alpha2,6-sialyltransferase
 RT family, ST6GalNAc III and IV.";
 RL J. Biol. Chem. 274:11958-11967(1999).
 RP [2]
 RP SEQUENCE OF 299-360 FROM N.A.
 RC STRAIN=CS7BL/6;
 RX MEDLINE=20198252; PubMed=10731711;
 RA Takashima S., Kurosawa N., Tachida Y., Inoue M., Tsuji S.;
 RT "Comparative analysis of the genomic structures and promoter
 RT activities of mouse Staa2,3Galb1,3GalNAc GalNAc2,6-sialyltransferase
 RT genes (ST6GalNAc III and IV): Characterization of their Sp1 binding.";
 RL J. Biochem. 127:399-409(2000).
 CC -1- FUNCTION: Involved in the biosynthesis of ganglioside GD1a from
 CC GM1b. Transfers CMP-NeuAc with an alpha-2,6-linkage to GalNAc
 CC residue on NeuAc-alpha-2,3-Gal-beta-1,3-GalNAc of glycoproteins
 CC and glycolipids. Prefers glycoproteins to glycolipids.
 CC -1- CATALYTIC ACTIVITY: CMP-N-acetylneuraminic acid + N-acetyl-alpha-
 CC galactosaminyl-(1-2>3)-beta-D-galactosyl-(1->3)-N-acetyl-D-
 CC galactosaminyl-R = CMP + N-acetyl-alpha-neuraminyl-(2->3)-beta-D-
 CC galactosaminyl-R.
 CC -1- PATHWAY: Glycosylation.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (By
 CC similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Comment=Experimental confirmation may be lacking for some
 CC isoforms;
 CC Name=1; Synonyms=Long 2;
 CC IsoId=Q9R2B6-1; Sequence=Displayed;
 CC Name=2; Synonyms=Long 1;
 CC IsoId=Q9R2B6-2; Sequence=VSP_001788;
 CC Name=3; Synonyms=Short;
 CC IsoId=Q9R2B6-3; Sequence=VSP_001787;
 CC -1- TISSUE SPECIFICITY: High expression in brain and colon and to a
 CC lesser extent in lung, heart, kidney, spleen and thymus.
 CC -1- DEVELOPMENTAL STAGE: Developmentally regulated.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase family 29.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: Y15780; CAB43514.1; -;
 CC EMBL: Y15780; CAB43515.1; -;
 CC EMBL: Y15779; CAB43507.1; -;
 CC EMBL: Y15779; CAB43508.1; -;
 CC EMBL: AJ007310; CAB07446.1; -;
 CC EMBL: Y19057; CAB93948.1; -;
 CC MGD: MGI:134184; Stac7d.
 CC InterPro: IPR001675; Glyco_trans_29.
 CC Pfam: PF00777; Glyco_transf_29; I.
 CC Alternative splicing; Glycoprotein; Glycosyltransferase; Golgi stack;

KM Signal-anchor; Transferase; Transmembrane.
 FT DOMAIN 1 71 Cytoplasmic (Potential).
 FT TRANSMEM 72 94 Signal-anchor for type II membrane
 FT protein (Potential).
 FT DOMAIN 95 360 Luminal, catalytic (Potential).
 FT DISULFID 134 283 By similarity.
 FT CARBOHYD 193 193 N-linked (GlcNAc...) (Potential).
 FT VARSPIC 1 58 Missing (in isoform 3).
 FT VARSPIC 1 28 /FTId=VSP_001787.
 FT VARSPIC 1 28 /FTId=VSP_001788.
 FT PORTINSP (in isoform 2).
 FT SEQUENCE 360 AA; 40773 MW; 7CF4101B8FPA369A CRC64;
 SO
 Query Match 88.5%; Score 1439; DB 1; Length 360;
 Best Local Similarity 88.1%; Pred. No. 1.7e-122;
 Matches 266; Conservative 11; Mismatches 25; Indels 0; Gaps 0;
 QY 1 MKAPGRVLVILICSVSAVYILLCCWAGPLCLATCLDHHPTGSRPTVGPPLHFGYS 60
 DB 59 MKAPGRLLTLTLCLTFSAVCVFLCCACPLCLATLDLHLPAPASTVPGPLHFGYS 118
 QY 61 SVDPGKPLVEPCRSACAVSSGQMLSGGLGABIDSAECVFRNQAPTVGFADVGQST 120
 DB 119 SVDPGKPLIHELCHSCAVNSGQMLSGGLGADAEVCLRMNQAPTVGFADVGQRTT 178
 QY 121 LRVVSHTSVPLLRNLSHTYQKARDITYMVGQGRHMDRYLGGRTTTLQLTRMYGLQ 180
 DB 179 LRVVSHTSVPLLRNLSHTYQKARDITYMVGQGRHMDRYLGGRTTTLQLTRMYGLQ 238
 QY 181 VYFTEEMMAVYCOIIPDEGKRNROSGFSLSGFMFLALCEIIVYVGMVSDYCR 240
 DB 239 VYFTEEMMAVYCOIIPDEGKRNROSGFSLSGFMFLALCEIIVYVGMVSDYCS 298
 QY 241 EKSHPSVPHYFPEKGRUDECOMYLHBOAPRSARHPTTEKAVSRMAKGRPIVAPHSWR 300
 DB 299 EKSPRSVPHYFPEKGRUDECOMYLHBOAPRSARHPTTEKAVSRMAKGRPIVAPHSWR 358
 QY 301 TE 302
 DB 359 AK 360
 RESULT 12
 Q6ZX20 PRELIMINARY; PRT; 302 AA.
 ID Q6ZX20;
 AC Q6ZX20;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE Alpha-2,6-sialyltransferase ST6GalNAc IV (fragment).
 GN Name=Stac7d;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar;
 RA Harduin-Lepera A., Martinez-Duncker I., Mollicone R., Delannoy P.,
 RA Oriol R.;
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ546871; CAG26700.1; -;
 DR GO: GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
 DR InterPro: IPR001675; Glyco_trans_29.
 DR Pfam: PF00777; Glyco_transf_29; I.
 KM Glycosyltransferase; Transferase.
 FT NON_TER 1 1
 FT NON_TER 1 1
 FT SEQUENCE 302 AA; 34332 MW; FA3F65E9B0C25886 CRC64;
 SO
 Query Match 88.3%; Score 1436; DB 2; Length 302;
 Best Local Similarity 88.1%; Pred. No. 2.6e-122;

OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Harduin-Lepers A., Martinez-Duncker I., Moillone R., Delannoy P.,
RA Oriol R.;
RT "Phylogeny of silyltransferases.",
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ620653; CAF05852.1 -
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 289 AA; 32780 MW; 9FB64C63BA5CA74 CRC64;

Query Match	64.4%	Score 1047.5;	DB 2;	Length 289;
Best Local Similarity	66.3%;	Pred. No. 6e-87;		
Matches 195;	Conservative 33;	Mismatches 55;	Indels 11;	Gaps 2;

```

QY 6 LVLVLLLSVVSAYVLLCCWAGRLPLCATCLIDHHPGSRPTVGPGLPFSGYSSVPG 65
Db 4 RVFVVLCCAAALSVLYVLCREA-----AGQD-----GSAATPAALSLQGYSRVPG 52
QY 66 KPLVNEPCRSCAVSSSGQMLGSGLGAELIDSAECVFRNQAPTVGFVADVGQRSTLRVS 125S
Db 53 KPLRAPRCRCVAVSSSGQMLGSHLGREIDGECVLRMNAHPAGEBEDVGTSIVRVS 112S
QY 126 HTSVPLLLRNYSHYFOKADLTLYMWGQGRHMDRVLGSRTYPTLLDTLTRYPCGLQVYFT 165S
Db 113 HTSVPLLRNQYVFPQOSKDTIYVLYGSPRSKRSREKGFPHRLMLKRLKEMPIQLYTLT 172S
QY 186 ERMAVYCDQIFQDETGGKNNRQSGSFLSTGMFTWILLALBCEELIVVYGWVDSYCREKSHP 245S
Db 173 EEKMAVCDQVFNQETGKNNLKSGSFLSTGMFTWILLAMELCENICVFGWVDSYCREKSHS 232S
QY 246 SVPHYVFEKGRLLDECOMYLAHQAQPSARPRITREKAVPSRWAKKRPIVFAHPQM 299
Db 233 SVPHYVFEKGRLLDECOMYLVHEBAQPAQHRFTIEKALFSSWAKKKOILIFSHPM 286

```

Search completed: December 10, 2004, 18:25:12
Job time : 194 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2004, 18:05:03 ; Search time 40 Seconds
(without alignments)
726.437 Million cell updates/sec

Title: US-10-019-114A-7

Sequence: 1 MKRPGRLVLLICSVFSAV.....FSRWAKRPVFAHPSWRT 302

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR 79: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	239	14.7	350	2	B54420
2	230	14.1	350	2	JC5251
3	228.5	14.1	350	2	A54420
4	221.5	13.6	524	2	JC8017
5	209.5	12.9	337	2	S36824
6	206.5	12.7	342	2	S55675
7	202.5	12.5	380	2	JC6321
8	200.5	12.3	359	2	JC4224
9	198.5	12.2	340	2	I54229
10	196.5	12.1	359	2	S52425
11	191	11.7	343	2	A45073
12	190.5	11.7	380	2	A56950
13	188.5	11.6	359	2	I59403
14	182.5	11.2	413	2	S41114
15	178	10.9	566	2	A49880
16	172	10.6	374	2	A45074
17	172	10.6	375	2	JN0618
18	168	10.3	406	2	A41734
19	164.5	10.1	356	2	A54032
20	164.5	10.1	526	2	JC7248
21	158	9.7	403	2	A28451
22	157	9.7	359	2	JC4973
23	155.5	9.6	342	2	JC5195
24	153.5	9.4	425	2	T00720
25	147	9.0	398	2	C86216
26	144.5	8.9	375	2	I48686
27	143	8.8	329	2	A48715
28	143	8.8	332	2	A49879
29	141.5	8.7	375	2	A46727

30	140.5	8.6	376	2	JC5600	alpha-N-acetylneur
31	139.5	8.6	375	2	I39169	siatyltransferase
32	137.5	8.5	404	2	A504871	Gal beta-1, 3GalNA
33	134	8.2	387	2	JB0364	lactosylceramide a
34	127.5	7.8	222	2	B54898	STX protein - huma
35	126	7.7	470	2	T49278	hypothetical prote
36	97	6.0	255	2	T52352	hypothetical prote
37	93.5	5.8	990	2	D83706	antibiotic merzsa
38	89	5.5	431	2	B37802	ctx protein - Erw
39	88.5	5.4	883	2	A96662	hypothetical prote
40	88	5.4	480	2	S14394	bindin precursor -
41	86.5	5.3	517	2	S32169	hypothetical prote
42	86	5.3	618	2	I48914	protein-tyrosine k
43	85.5	5.3	461	2	B41313	probable transposa
44	84	5.2	322	2	S66766	probable membrane
45	84	5.2	642	1	S52111	utromodulin precurs

ALIGNMENTS

```
RESULT 1
B54420
beta-galactoside alpha-2,3-sialyltransferase (EC 2.4.99.4) ST3GAL-2 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 09-Jul-2004
C:Accession: B54420
R:Lee, Y.C.; Kojima, N.; Wada, E.; Kurosawa, N.; Nakacka, T.; Hamamoto, T.; Tsuji, S.
J. Biol. Chem. 269, 10028-10033, 1994
A:Title: Cloning and expression of cDNA for a new type of Galbeta1,6GalNAc alpha2,3-sialyl
A:Cross-references: A54420; MUID:94193584; PMID:8144500
A:Reference number: A54420
A:Accession: B54420
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-350 <LEE>
C:Superfamily: beta-galactoside alpha-2,3-sialyltransferase
C:Keywords: glycosyltransferase; transmembrane protein

Query Match 14.7%; Score 239; DB 2; Length 350;
Best Local Similarity 34.3%; Pred. No. 2,2e-13;
Matches 71; Conservative 24; Mismatches 80; Indels 32; Gaps 8;

QY 59 YSSVPDQKPL-VREP--CRCAVYSSGOMLGGSLGAEIDSABEVFMNQAPTYGFADV 115
DB 132 FOYVGENPYFRDPQCCRCVAVGNSGNLRGSGYGVDSHNFIMMNOAPVGFEDV 191
QY 116 GORSTLRVSGHTSVPLLRVYSHYFQ--KARDLYWV-----WGQGRMDRVLGRTYRTL 169
DB 192 GSRTHHFMTPEAKNULPANVSPVLVPRKALDLMWIMASLSTQDIR-----TAPV 243
QY 170 LQLRMYPGIQQVYTFERRMAVYCDQIFQDETGNRRQSGFLSTGWTMLALECEIV 229
DB 244 KSFLLRV-----DKEKQVQIYNPAFPKYLIDRWTEHHGRVPSGMLVLFALHVCDEVN 295
QY 230 VYGVNDSYCREKSHPSVPYHYPEKGR 256
DB 296 VYGFAGDS--RGWNR-----HYWENNR 315

RESULT 2
JC5251
beta-galactoside alpha-2,3-sialyltransferase (EC 2.4.99.4) - human
C:Species: Homo sapiens (man)
C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: JC5251; MUID:97079181; PMID:8920913
R:Kim, Y.T.; Kim, K.S.; Kim, S.H.; Kim, C.H.; Ko, J.H.; Choe, I.S.; Tsuji, S.; Lee, Y.C.
Biochem. Biophys. Res. Commun. 228, 324-327, 1996
A:Title: Molecular cloning and expression of human Gal beta 1,3GalNAc alpha 2,3-sialyltr
A:Reference number: JC5251; MUID:97079181; PMID:8920913
A:Accession: JC5251
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
```

A;Residues: 1-350 <KIM>
 A;Cross-references: UNIPROT:Q16842; GB:U63090; NID:G1773282; PIDN:AB40389.1; PID:G17732
 A;Experimental source: liver
 R;Giordanengo, V.
 Submitted to the EMBL Data Library, March 1996
 A;Reference number: H00561
 A;Accession: G01021
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-350 <GIO>
 A;Cross-references: EMBL:X96667; NID:G1235530
 C;Comment: This enzyme catalyzes the transfer of sialic acid from CMP-NeuAc to the terminal
 substrate preference for glycolipid than for O-linked oligosaccharides of glycoproteins.
 C;Genetics:
 A;Gene: ST3 (0) -II
 C;Superfamily: beta-galactoside alpha-2,3-sialyltransferase
 C;Keywords: glycosyltransferase

Query Match 14.1%; Score 230; DB 2; Length 350;
 Best Local Similarity 33.3%; Pred. No. 1,3e-12;
 Matches 69; Conservative 24; Mismatches 82; Indels 32; Gaps 8;

QY 59 YSSVDPGKPL-VREP--CRSCAVSSSGQMLGSLGAEDISAECPFRMNOAPTVGFADY 115
 DB 132 FOIVPGENPYRFRDPHCRCACAVVGNLRSYGVDVGDHNFMMNOAPTVGFADY 191
 QY 116 GQSTLRVSHTSVPLLRVSHYFO--KARDTLVWV---WGGRHMDVLLGGRTYRTL 169
 DB 192 GSRTHHFMPESAKNLPANVSFVLVFPKVLDMWLSALSTGQIRF-----TYAPV 243
 QY 170 LQITRMYPGLQVYFTFRMAAYCDQIFQDETGNKRSGSFLSTGWTMTLALCEIY 229
 DB 244 KSFRLV-----DKKKQIYNPAFFKYIHDRWTEHHGRVPTGMLVLFALHVCDEV 255
 QY 230 VYGVNDSDYCREKSHPSVRYHYEKGK 256
 DB 296 VYGFAGDS--RGNWV-----HYWENNR 315

RESULT 3
 A54420
 beta-galactoside alpha-2,3-sialyltransferase (EC 2.4.99.4) ST3GAL1.2 - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 09-Jul-2004
 C;Accession: A54420
 R;Lee, Y.C.; Kojima, N.; Wada, E.; Kurosawa, N.; Nakaoka, T.; Hamamoto, T.; Tsuji, S.
 J. Biol. Chem. 269, 10028-10033, 1994
 A;Title: Cloning and expression of cDNA for a new type of Galbeta1,6Galnac alpha2,3-sial
 A;Reference number: A54420; MUID:94193584; PMID:8144500
 A;Accession: A54420
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-350 <LEE>
 A;Cross-references: UNIPROT:Q11204; GB:X76989; NID:G475203; PIDN:CAA54294.1; PID:G475204
 C;Superfamily: beta-galactoside alpha-2,3-sialyltransferase
 C;Keywords: glycosyltransferase; transmembrane protein

Query Match 14.1%; Score 228.5; DB 2; Length 350;
 Best Local Similarity 34.6%; Pred. No. 1.8e-12;
 Matches 65; Conservative 23; Mismatches 75; Indels 25; Gaps 6;

QY 59 YSSVDPGKPL-VREP--CRSCAVSSSGQMLGSLGAEDISAECPFRMNOAPTVGFADY 115
 DB 132 FOIVPGENPYRFRDPHCRCACAVVGNLRSYGVDVGDHNFMMNOAPTVGFADY 191
 QY 116 GQSTLRVSHTSVPLLRVSHYFO--KARDTLVWV---WGGRHMDVLLGGRTYRTL 169
 DB 192 GSRTHHFMPESAKNLPANVSFVLVFPKVLDMWLSALSTGQIRF-----TYAPV 243
 QY 170 LQITRMYPGLQVYFTFRMAAYCDQIFQDETGNKRSGSFLSTGWTMTLALCEIY 229
 DB 244 KSFRLV-----DKKKQIYNPAFFKYIHDRWTEHHGRVPTGMLVLFALHVCDEV 255

QY 230 VYGVNDSDS 237
 DB 296 VYGFAGDS 303

RESULT 4
 JC8017
 beta-galactoside alpha-2,6-sialyltransferase (EC 2.4.99.1) II - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 04-Apr-2004 #sequence_revision 04-Apr-2004 #text_change 04-Apr-2004
 C;Accession: JC8017
 R;Takashima, S.; Tsuji, S.; Tsujimoto, M.
 J. Biochem. 134, 287-296, 2003
 A;Title: Comparison of the enzymatic properties of mouse beta-galactoside alpha2,6-sialyl
 A;Reference number: JC8017; PMID: 12966079
 A;Accession: JC8017
 A;Molecule type: mRNA
 A;Residues: 1-524 <TKA>
 A;Cross-references: DDBJ:AB095093
 C;Comment: This enzyme is an oligosaccharide-specific enzyme and a type II transmembrane
 cid to galactose of the Gal beta1,4 GlcNAc structure through an alpha 2,6-linkage.
 C;Genetics:
 A;Gene: ST6GAL1
 A;Map position: 17C
 A;Intons: 310/1; 342/3; 376/3; 435/1
 C;Keywords: oligosaccharide-specific; sialyl motif; sialyltransferase; ST6GAL1; transmen

Query Match 13.6%; Score 221.5; DB 2; Length 524;
 Best Local Similarity 28.6%; Pred. No. 1.2e-11;
 Matches 64; Conservative 38; Mismatches 83; Indels 39; Gaps 8;

QY 45 GSRPTVPG-----PLHPSGYSSVDPGKPLVREPGRSCAVSSSGQMLGSLGAEDISAEV 100
 DB 270 GMRPLVPGVPLSGJHRLGS-----SCAVVMSKAILNLSIGEIDSHDAV 315
 QY 101 FRMNOAPTVGFADVQGRSTLRVSHTSVPLLRVSHYFOKA--RDTLYWVGGRHM 157
 DB 316 LRFNSAPTRGYEDVGNKTTVRIINSQ---ILANSHHFIDSLYKVDVLLVMDAPYS 371
 QY 158 DRV-----LGGTYRTLQITRMYPGLQVYFTFRMAAYCDQIFQDETGNKRSGSFL 211
 DB 372 ANLNLWKKPDYNLFPFYIQRKRYTPQPIYLPFIQWLDIENREKTIQNP-- 429
 QY 212 STGMFTMTLALCEIIVVYGVNDSDYCREKSHPSVYH--YPE 253
 DB 430 SSGFGLIMSMCKEVHYEYIPSVRQIELCH---YHELYVD 469

RESULT 5
 S36824
 beta-galactoside alpha-2,3-sialyltransferase (EC 2.4.99.4) - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 22-Jan-1994 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 C;Accession: S36824
 R;Lee, Y.C.; Kurosawa, N.; Hamamoto, T.; Nakaoka, T.; Tsuji, S.
 Eur. J. Biochem. 216, 377-385, 1993
 A;Title: Molecular cloning and expression of Gal-beta1,3GalNac-alpha-2,3-sialyltransfer
 A;Reference number: S36824; MUID:93387288; PMID:8375377
 A;Accession: S36824
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-337 <LEE>
 A;Cross-references: UNIPROT:P54751; EMBL:X73523; NID:G402214; PIDN:CAA51919.1; PID:G40221
 C;Superfamily: beta-galactoside alpha-2,3-sialyltransferase
 C;Keywords: glycosyltransferase

Query Match 12.9%; Score 209.5; DB 2; Length 337;
 Best Local Similarity 32.6%; Pred. No. 8.3e-11;
 Matches 61; Conservative 32; Mismatches 65; Indels 29; Gaps 7;

QY 73 CRSCAVSSSGQMLGSLGAEDISAECPFRMNOAPTVGFADVQGRSTLRVSHTSVPL 132
 DB 136 CRCAVAVVGNLKDSSYGEIDSHDFVLRMNAKAPTVGFADVSGRTTHLV----- 187

Query Match 12.7%; Score 206.5; DB 2; Length 342;
 Best Local Similarity 28.1%; Pred. No. 1.5e-10;
 Matches 61; Conservative 37; Mismatches 68; Indels 51; Gaps 7;

133 LRVNHYFOKARDTLVWVGQGRHMDR--VLGGRTYTLTLQTRMY-----PGLQVTFTE 186
 188 ----YPSFRRLGNGVNVLPFKTDLQWISATTTCT---ITHTYVPVPRKIKV--QE 239

187 RMAYCDQIFQDETGNKRRSGSFLSTGWTMLALCEIIVYGVNSDYCREKSHPS 246
 240 KILIHAFKIVYFDWMLQGHGPRSTGLISIFSIHCDEVDLYGFGADS--KGNWH-- 295

247 VPHYFE 253
 296 ---HYME 299

RESULT 6

Gal-beta-1,3galNac alpha-2,3-sialyltransferase - chicken
 C/Species: Gallus gallus (chicken)
 C/Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
 C/Accession: S55675

R;Kurosawa, N.; Hamamoto, T.; Inoue, M.; Tsuji, S.
 Biochim. Biophys. Acta 1244, 216-222, 1995

A;Title: Molecular cloning and expression of chick Gal-beta-1,3galNac alpha-2,3-sialyltr
 A;Reference number: S55675; PMID:95284088; PMID:7766661

A;Accession: S55675

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-342 <KUR>

A;Cross-references: UNIPROT:Q11200; GB:X80503; NID:g975654; PIDN:CAA56666.1; PID:g975655

C;Superfamily: beta-galactoside alpha-2,3-sialyltransferase

Query Match 12.7%; Score 206.5; DB 2; Length 342;
 Best Local Similarity 28.1%; Pred. No. 1.5e-10;
 Matches 61; Conservative 37; Mismatches 68; Indels 51; Gaps 7;

50 VPG---PLHPSGYSVPDGRPLVREPCRSQAVSSSGOMLGSGLGAEIDSACVFRNQA 106
 126 IPDRDPLQGRGFRS-----CRCAVVGNSGNLRSQYQGDIDSHDFLARWRA 174

107 PTMGFEADVQGRSTLRVSHTSVPLLRNYSHTFQKARDTLVWVGQGRHMDRYLGRRTY 166
 175 PTIGESDVGSKTHHFV-----YPSYKEIAENVSMT-----VIPPKTL 214

167 RTLLQLTRMYPGIQLVTFE-----ERMAYCCQIFQDETGNKRRSGSFLSTGWF 216
 215 DLRWIVTALTGTGTFYVVPVRKIKYRKEKVLVLPSPKTYVENWLVQNHGRYPSTGL 274

217 TMLALCEIIVYGVNSDYCREKSHPSVPHYFE 253
 275 SVTFALHVCDEVNVYGFADS---KGN---WHHYME 304

RESULT 7

JC6321
 CMP-NeuAc-GM3 alpha2-8 sialyltransferase (EC 2.4.99.-) [imported] - rat
 N;Alternate names: GD3-synthase; GM3-synthase
 C;Species: Rattus norvegicus (Norway rat)
 C/Date: 14-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
 C/Accession: JC6321

R;Zeng, G.; Gao, L.; Yu, R.K.
 Gene 187, 131-134, 1997

A;Title: Cloning of the cDNA coding for rat brain CMP-NeuAc:GD3 alpha2-8 sialyltransferase
 A;Reference number: JC6321; PMID:97225806; PMID:9073076

A;Accession: JC6321

A;Molecule type: mRNA

A;Residues: 1-380 <ZEN>

A;Cross-references: UNIPROT:P97877; GB:U55938

A;Experimental source: brain

C;Comment: This enzyme is one of the sialyltransferases and adds one more alpha-2,8-link
 C;Superfamily: alpha-N-acetylnneuraminate alpha-2,8-sialyltransferase
 C;Keywords: glycosyltransferase
 F;18-33/Domain: transmembrane #status predicted <TM>

Query Match 12.5%; Score 202.5; DB 2; Length 380;
 Best Local Similarity 30.4%; Pred. No. 4e-10;
 Matches 66; Conservative 29; Mismatches 79; Indels 43; Gaps 8;

60 SSVPDGRPLVREPCRSQAVSSSGOMLGSGLGAEIDSACVFRNQAFTVGFADVGRS 119
 146 SLPLDVSPIILNKRYNCAVVGNSGILTSGCGGEIDKSDPVFCNPAFTAFAFKDYDKT 205

120 TLAVGHTSVPLLRNYSHTFQKARDTL-----YMWVGQGRHMDRYLGRRTYRT 168
 206 NLT---TENPSILEKTYNNLTLTDQNNFFLSKLQDAILM-----IPAFFHT 252

169 ILQTRMYPGIQLVTFE-----MMAYCDQIFQ--DETGNKRRSGSFLSTGWTML 220
 253 SATVTRTL-----VDFVEHKGQKVLQVAMPNGIMGVNRYWKQKHLSPKRLSTGFLMYTL 308

221 ALFLCEIIVYGVNS---DSYCREKSHPSVPHYFEK 254
 309 ASAICEIHLYGFWPGFDPNTR---DLFVHYDK 341

RESULT 8

JC4224
 alpha-N-acetylnneuraminate alpha-2,8-sialyltransferase (EC 2.4.99.8) - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 19-Oct-1995 #sequence_revision 08-Feb-1996 #text_change 09-Jul-2004
 C/Accession: JC4224

R;Yoshida, Y.; Kojima, N.; Tsuji, S.
 J. Biochem. 118, 658-664, 1995

A;Title: Molecular cloning and characterization of a third type of N-glycan alpha 2,8-si
 A;Reference number: JC4224; PMID:96115941; PMID:8690732

A;Accession: JC4224

A;Molecule type: mRNA

A;Residues: 1-359 <YOS>

A;Cross-references: UNIPROT:Q64692; EMBL:X86000; NID:g2665331; PIDN:CAA59992.1; PID:g122

A;Experimental source: lung

C;Comment: This enzyme exhibits activity specific toward sialylated glycoproteins. This e
 C;Superfamily: alpha-N-acetylnneuraminate alpha-2,8-sialyltransferase
 C;Keywords: glycoprotein; glycosyltransferase; lung; transmembrane protein
 F;8-20/Domain: transmembrane #status predicted <TM>
 F;141-185/Region: sialyl signature L
 F;280-302/Region: sialyl signature S
 F;50,74,119,204,219/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 12.3%; Score 200.5; DB 2; Length 359;
 Best Local Similarity 27.9%; Pred. No. 5.5e-10;
 Matches 70; Conservative 38; Mismatches 98; Indels 45; Gaps 9;

59 YSVPDGRPLVREPCRSQAVSSSGOMLGSGLGAEIDSACVFRNQAFTVGFADVGR 118
 125 HSLLPEVSPKRNRRPRTCAVVGNSGILLDSGCKEIDSHNPFVLRCLAPVVEFADVGR 184

119 SLRVVSHTSVPLLRNYS-----HYFOKARDTL-----YMWVGQGRHMDRYLGR 164
 185 SDPTTNMPSVVGAFAGFGRNEDREKFRHLSTMLNDSVLMIIPAFVYVGGEKHEWV--NAL 243

165 TYRTLLQTRMYPGIQLVTFE-----YTTERMMAYCDQIFQDETGNKRRSGSFLSTGFTM 218
 244 ILNKKLQVTRAYPSLRLIHAVRGYWLTKVPI-----KRP-----STGLMY 265

219 ILALCEIIVYGVNSDYCREKSHPSVPHYFEKGRLDCEQMYLAHQAPSAHFIT 278
 286 TLATRCDEIHLXGFW--PPKDLNKAIVKTHYD---DLKTRYFSNAPSHPMPLEFKT 339

279 EKAVFSRMACK 289
 340 LNVTLNHRGALK 350

RESULT 9

beta-galactoside alpha-2,3-sialyltransferase (EC 2.4.99.4) - human
 C/Species: Homo sapiens (man)

Query Match	11.7%	Score 190.5	DB 2	Length 380
Best Local Similarity	30.0%	Pred. No. 4.5e-09		
Matches	65	Conservative 28	Mismatches 81	Indels 43
				Gaps 8

Qy	60	SSVPDGKPLVREPCRS	CAVSSSGOMLGSGLA	EIDSAECVFRMNQAPTVGEADVGRS	119
Db	146	SILPDVSEINMKRNV	CAVNGNSGILTSG	OCGEIDKSDPFSRCNFAPTFAFHDDVDRKT	205
Qy	120	TLRVASHSVGLLRN	SNHYFQKARDTL	-----YVWQGRHMDVLAGRTYRT	168
Db	206	NLT-----TFNDS	ILEKTYNNLLIT	IODRNNFLSLKLDGAILM-----IAFFPHT	252
Qy	169	LLQLTRMYGLQVY	TFETR-----MMA	YCDQIFQ--DETGNRRQSGSFLSTGWFMTIL	220
Db	253	SATVTRTL-----	VDFFVEH	RGOLKVQLAMPNINQMHWVRVYKKNLSPKRLSTGILMYTL	308
Qy	221	ALAEICEELIVVGM	VS-----DSYCREK	SHSVVYHYPEK	254
Db	309	ASAICEEIHLVGF	WPFQFPDPTRE	----DLYPHYIDK	341

```

RESULT 13
159403
alpha-2,8-poly(sialyl)transferase - human
C:Species: Homo sapiens (man)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 09-Jul-2004
C:Accession: 159403
R:Nakayama, J.; Fukuda, M.N.; Predelte, B.; Ranscht, B.; Fukuda, M.
Proc. Natl. Acad. Sci. U.S.A. 92, 7031-7035, 1995
A>Title: Expression cloning of a human poly(sialyl)transferase that forms the poly(sialyl)
A:Reference number: 159403; MUID:95350205; PMID:7624364
A:Accession: 159403
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-359 <RBS>
A:Cross-references: UNIPROT:Q92187; GB:L41680; NID:9945220; PIDN:AAC41775.1; PID:99452221
C:Gene: GDB:PSR
C:Genetics:
A:Gene: GDB:PSR
C:Superfamily: alpha-N-acetylneuraminatase alpha-2,8-sialyltransferase

Query Match 11.6%; Score 188.5; DB 2; Length 359;
Beat Local Similarity 27.1%; Pred. No. 6.3e-09;
Matches 68; Conservative 39; Mismatches 99; Indels 45; Gaps 9;

QY 59 YSSVPDGKPLVREPCRCAYVSSGQMLGSLGLAIDSAECVFERNQAPTYGFADYQGR 118
Db 125 HSLPEVSPMKRRRPFKTCAYVGNSGIILDSCEGKEIDSHNFVIRCNLAPVVEFADYGTK 184
QY 119 STLKVVSHTSVPLD---LRNYS-----HYQKARDTL-----YMWGQGHMRYLQGR 164
Db 185 SDFITMNSVYQRAFGGFRNEDREKFPVHLISMLNDSYLTWIPAMVNVKGGEGHVEWV-NAL 243
QY 165 TYRITLQITRMYPGHVV-----YTFTERRMAAYCDQIFQDETGNKRQSGSFLSTGWFTM 218
Db 244 ILKAKLKVTRTAYPSRLRLHAVRGWTLYTKVYI-----KRP-----STGLMY 285
QY 219 ILALCELCEIVVYGVNDSYCRKSHPSVPYHYFEKGLDCEQWYLAHQAPRSANRFTT 278
Db 286 TLARFPCDEIHLHYGW--PFPKDLNGKAVKXHYD---DLKRYTFNSAPSHMRDLEPKT 339
QY 279 EKAVFSRMAYK 289
Db 340 LNVLHNRGALK 350

RESULT 14
S41114
Gal beta 1,4 GlcNAc alpha 2,6-sialyltransferase - chicken
C:Species: Gallus gallus (chicken)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S41114
R:Kurobawa, N.; Kawabeaki, M.; Hamamoto, T.; Nakaoaka, T.; Lee, Y.C.; Arita, M.; Teuji, S.

```

Eur. J. Biochem. 219, 375-381, 1994
 A:Title: Molecular cloning and expression of chick embryo Gal-beta-1,4GlcNAc-alpha-2,6-sialyltransferase
 A:Reference number: 581114; MUID:94139712; PMID:8307003
 A:Accession: S4114
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-413 <KID>
 A:Cross-references: UNIPROT:092182; EMBL:X75558; NID:G452750; PIDN:CA53235.1; PID:G45275
 C:Superfamily: beta-galactoside alpha-2,6-sialyltransferase

Query Match	11.2%	Score 182.5	DB 2	length 413
Best Local Similarity	26.1%	Pred. No. 2.5e+08		
Matches	61	Conservative	39	Mismatches 95; Indels 39; Gaps 8

Qy	52	GPLHFGSYSSVDPGCKPLVREPCR--SCAVNSSGCMATGSGCAETDSAECPFRNQAPLV	109
Db	164	GPFSSSSMOHYLPDKSINETVGRGRCVASSASGLSKSHLGPELDSDAVLRFNGAPYK	223
Qy	110	GFEADVQORSTPLRVVSHTSVPLLRNASHYFQKARDTLV-----MVMGCGRHMDV-----	160
Db	224	GFOEDVQCKTIRLVNSQGLVVEEQF-----LKDALYNGILLVMPDAPRYNAEIHMEY	277
Qy	161	-----LGRYRTLLQLTRMYRGLQVYTFERMAVYCDQIFODETEKKNRRSGFSL	212
Db	278	RKPDYKFEFAKSYRI-----RHDEQFYILNPKMQWOLMDILQENSLYENHPNP--S	329
Qy	213	TGWFETMLLAEIGELIVYGVMSDYSCEKSHSPRYVFERKGRLEDGOMYLAH	266
Db	330	SGMLGIVMMTLCEVDVY-----EFLPSKQOTDI-CHYUQKFNDACTMGATH	377

RESULT 15
A49880
alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase (EC 2.4.99.3) - chicken
C:Species: Gallus gallus (chicken)
C:Date: 30-Jun-1995 #sequence_revision 30-Jun-1995 #text_change 09-Jul-2004
C:Accession: A49880
R:Kurosawa, N.; Hamamoto, T.; Lee, Y.C.; Nakaoka, T.; Kojima, N.; Tsuji, S.
J. Biol. Chem. 269, 1402-1409, 1994
A:Title: Molecular cloning and expression of GalNAc alpha2,6-sialyltransferase.
A:Reference number: A49880; MUID:94117458; PMID:8288607
A:Accession: A49880
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-566 <KUR>
C:Cross-references: UNIPROT:092183; GB:X74946; NID:G453196; PIDN:CAA52902.1; PID:G453197
C:Superfamily: alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase.
A:Keywords: glycosyltransferase

	Query Match	10.9%; Score 178; DB 2; Length 566;
	Best Local Similarity	25.2%; Pred. No. 9.3e-08;
	Matches	Conservative 63; Mismatches 80; Indels 70; Gaps 9
Qy	73 CRCAVAVSSSGOMIGSLGAETDSACEYRMMNOAPTVGFEPADVGRGSTRLVVSHTSVPL	132
Dd	: : : : : : : : : : : : : :	
	337 CISCAVVANGNGILNNSGWGEIDSDHYDFRVSAVGVIKYEKDVGKTSTFFGYTAIVLVS	396
Qy	133 LRNTSHY-FOKARDPLLVMWGCGRHMDRLVLSGARFRTLQLTTRMPGLGV-	181
Dd	: : : : : : : : : : : : : : : : :	
	397 LQNIGHGGFKLI-----PQGHHI-----RYIHFLVAIRDVEWTKALLDLDIRKG	441
Qy	182 -----YTFERMAYCYDQIFODETGKKRRROGSGL--SICWM	215
Dd	: : : : : : : : : : : : : : : : :	
	442 FLNYTYGRPRERFEDEDFTMNKYTLVAHPDEL--RLYKRFLPKSNLKXPWMLYNPTTGA	498
Qy	216 PTMIALTELCEIIIVYGVDSDSYCRKSHPSEVPYYHPEKGRLLDECOMYLAAHQAPSARH	275
Dd	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :	
	499 LLILLTALTHLCORVASAGYTE----GHQXCSPHYDK-EMKRLVFVNHD-----	543
Qy	276 FITEKAVFSR	285
Dd	: : : : : : : : : : : : : :	
	544 FNLEKVAMWR	553

Tue Dec 14 09:16:19 2004

us-10-019-114a-7.rpt

Page 6

Search completed: December 10, 2004, 18:25:57
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 18:02:40 ; Search time 25 Seconds

(without alignments)
801.121 Million cell updates/sec

Title: US-10-019-114A-7

Perfect score: 1626

Sequence: 1 MKAPGRLVLIILCSVFSVSAV.....FSRMACKRPIVPAHPSWRIB 302

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1626	100.0	302 4 US-09-599-360B-75	Sequence 75, App1
2	639.5	39.3	336 3 US-09-334-601-13	Sequence 13, App1
3	494	30.4	95 4 US-09-621-976-3915	Sequence 3915, App1
4	198.5	12.2	340 2 US-08-446-875-16	Sequence 16, App1
5	196.5	12.1	359 1 US-08-503-133A-2	Sequence 2, App1
6	196.5	12.1	359 2 US-08-576-775A-2	Sequence 2, App1
7	196.5	12.1	359 3 US-08-972-448-2	Sequence 2, App1
8	196.5	12.1	359 4 US-08-999-545-2	Sequence 2, App1
9	191	11.7	343 2 US-08-446-875-2	Sequence 2, App1
10	191	11.7	343 3 US-08-102-385G-2	Sequence 2, App1
11	190.5	11.7	339 1 US-08-626-994A-3	Sequence 3, App1
12	190.5	11.7	339 2 US-08-957-742-3	Sequence 3, App1
13	190.5	11.7	364 1 US-08-626-994A-1	Sequence 1, App1
14	190.5	11.7	364 2 US-08-957-742-1	Sequence 1, App1
15	178	10.9	355 3 US-08-666-367B-6	Sequence 6, App1
16	178	10.9	355 4 US-09-143-438-6	Sequence 6, App1
17	178	10.9	566 2 US-08-666-367B-5	Sequence 5, App1
18	178	10.9	566 3 US-09-143-438-5	Sequence 5, App1
19	173.5	10.7	339 3 US-09-334-601-10	Sequence 10, App1
20	173.5	10.7	339 4 US-09-334-601-21	Sequence 21, App1
21	172	10.6	375 2 US-08-446-875-10	Sequence 10, App1
22	172	10.6	375 3 US-08-102-385G-10	Sequence 10, App1
23	171	10.5	376 2 US-08-666-367B-8	Sequence 8, App1
24	171	10.5	376 3 US-09-143-438-8	Sequence 8, App1
25	170	10.5	374 2 US-08-446-875-4	Sequence 4, App1
26	170	10.5	374 3 US-08-102-385G-4	Sequence 4, App1
27	168	10.3	90 3 US-09-334-601-20	Sequence 20, App1

28	168	10.3	406 1 US-08-446-777-4	Sequence 4, App1
29	168	10.3	767 1 US-08-446-777-6	Sequence 6, App1
30	168	10.3	767 2 US-08-446-777-8	Sequence 8, App1
31	166.5	10.2	90 3 US-09-334-601-22	Sequence 22, App1
32	166.5	10.2	90 4 US-09-334-601-24	Sequence 24, App1
33	165	10.1	600 4 US-10-140-002-462	Sequence 462, App1
34	163.5	10.1	356 4 US-08-361-304A-2	Sequence 2, App1
35	159	9.8	340 2 US-08-102-385G-18	Sequence 18, App1
36	157.5	9.7	90 3 US-09-334-601-23	Sequence 23, App1
37	148.5	9.1	90 3 US-09-334-601-25	Sequence 25, App1
38	145.5	8.9	90 3 US-09-334-601-26	Sequence 26, App1
39	143	8.8	329 1 US-07-991-587A-7	Sequence 7, App1
40	143	8.8	329 2 US-08-309-985-7	Sequence 7, App1
41	143	8.8	332 2 US-08-446-875-12	Sequence 12, App1
42	143	8.8	332 3 US-08-102-385G-12	Sequence 12, App1
43	143	8.8	384 3 US-09-334-601-8	Sequence 8, App1
44	142.5	8.8	77 2 US-08-102-385G-29	Sequence 29, App1
45	142	8.7	55 2 US-08-102-385G-15	Sequence 15, App1

ALIGNMENTS

RESULT 1
US-09-599-360B-75
; Sequence 75, Application US/09599360B
; Patent No. 6548633
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bouguerelet, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: GENEST. 0500P3
; CURRENT APPLICATION NUMBER: US/09/599,360B
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: 60/113,686
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/141,032
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/469,099
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patent.pm
; SEQ ID NO 75
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -18..-1
US-09-599-360B-75
Query Match 100.0%; Score 1626; DB 4; Length 302;
Best Local Similarity 100.0%; Pred No. 3.4e-170; Indels 0; Gaps 0;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKAPGRLVLIILCSVFSVSAVYILLCWAGLPCLATCLDHHFPTGSRPTVPGPLHFSYGS 60
Db 1 MKAPGRLVLIILCSVFSVSAVYILLCWAGLPCLATCLDHHFPTGSRPTVPGPLHFSYGS 60
QY 61 SVPDGRLVREPRSCAVVSSSGMTGSGIGAEIDBAECVFRNQAPTYGFELDVQGRST 120
Db 61 SVPDGRLVREPRSCAVVSSSGMTGSGIGAEIDBAECVFRNQAPTYGFELDVQGRST 120
QY 121 LRVVSHTSVPLLRNYSYFQKARDTLVYWGGRHMDVLAGRTVTRTLQLTRMTPLGLQ 180
Db 121 LRVVSHTSVPLLRNYSYFQKARDTLVYWGGRHMDVLAGRTVTRTLQLTRMTPLGLQ 180
QY 181 VYTFERRMAVCDQIFQDETGNRRSGSPLSTGFTMTLALCEIIVYGVNSDYCR 240
Db 181 VYTFERRMAVCDQIFQDETGNRRSGSPLSTGFTMTLALCEIIVYGVNSDYCR 240
QY 241 EKSHPSVPPIHYEKGKLDCEQMTLAHQAPRSARHRTTEKAVPSRAKXRPVFAHPSMR 300
Db 241 EKSHPSVPPIHYEKGKLDCEQMTLAHQAPRSARHRTTEKAVPSRAKXRPVFAHPSMR 300

Db 241 EKSHPSVPIYFEKGLDECOMYLAHQAPRSARHPTTEKAVSRWAKRPVFAHPSWR 300
QY 301 TE 302
Db 301 TE 302

RESULT 2

US-09-334-601-13
Sequence 13, Application US/09334601
Patent No. 6280989
GENERAL INFORMATION:
APPLICANT: Kapitonov, Dmitri
APPLICANT: Yu, Robert
TITLE OF INVENTION: NOVEL SIALYLTRANSFERASES
FILE REFERENCE: VCUIP-6
CURRENT APPLICATION NUMBER: US/09/334,601
CURRENT FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 94
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 336
TYPE: PRT
ORGANISM: Homo sapiens
US-09-334-601-13

Query Match 39.3%; Score 639.5; DB 3; Length 336;
Best Local Similarity 47.9%; Pred. No. 1,1e-61;
Matches 125; Conservative 44; Mismatches 89; Indels 3; Gaps 3;

QY 44 TGSRPVTP-GEHFSGYSPVDGKPLVREPCSCAVVSSSGQMLGSGIGAIDSACVFR 102
Db 64 TQGRPVAGRPIDGVLGVADHKPL-KMCRDCAVTVSSGHLHSRQSGIDQTECVIR 122
QY 103 MNQAPVGFENDVQGRSTLRVSHTSVPLLRNYSHFQKARDTLVWVGGRMDRVLG 162
Db 123 MNDPRTGYGRDVQNRSLRYLHSSIORILKNHDLINVSQGVTFIFWGPSSYMRDCK 182
QY 163 GRVYRTLLQLTRMYPGLOVYTFTERMAVCDQIFQDETGNKRRSGSFLSTGWTMLAL 222
Db 183 GQVNNHLLSQVPLKAFMITRHKMLQFDELFKQETGKORKISNTWLSTGWTMTIAL 242
QY 223 ELCEEIVVYGWSSGYCEKSHSPVPIYFEKGLDECOMYLAHQAPR-SARHPTTEKA 281
Db 243 ELCDRIIVYGWPPDFCDPNHPSVPPIHYEPPGPDCEMTYLSHERGKSGSHRPTTEKR 302
QY 282 VFSRWAKRPVFAHPSWRT 302
Db 303 VFKWARTFNIHFQPDWKPE 323

RESULT 3

US-09-621-976-3915
Sequence 3915, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 3915
LENGTH: 95
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -18...-1

NAME/KEY: UNSURE
LOCATION: 59
OTHER INFORMATION: Xaa = Ala,Thr
NAME/KEY: UNSURE
LOCATION: 56
OTHER INFORMATION: Xaa = Arg,Ser
NAME/KEY: UNSURE
LOCATION: 54
OTHER INFORMATION: Xaa = His,Pro
US-09-621-976-3915

Query Match 30.4%; Score 494; DB 4; Length 95;
Best Local Similarity 96.8%; Pred. No. 1.7e-46;
Matches 92; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRAPEGLVILICSVFSAVYILLCWAGLPCLATCDLDPHTGSRPTVPGPLHPSGY 60
Db 1 MRAPEGLVILICSVFSAVYILLCWAGLPCLATCDLDPHTGSRPTVPGPLHPSGY 60
QY 61 SVPDGKPLVREPCSCAVVSSSGQMLGSGIGAIDSACVFR 95
Db 61 SVPDGKPLVREPCSCAVVSSSGQMLGSGIGAIDSACVFR 95

RESULT 4

US-08-446-875-16
Sequence 16, Application US/08446875
Patent No. 5858751

GENERAL INFORMATION:
APPLICANT: Paulson, James C.
APPLICANT: Wen, Xiaohong
APPLICANT: Livingston, Brian Duane
APPLICANT: Gillespie, William
APPLICANT: Keim, Sorge
APPLICANT: Burlingame, Alma L.
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Identification and Synthesis of Sialyltransferases
NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:
ADDRESSER: Poms, Smith, Lande & Rose
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: CA
COUNTRY: USA

ZIP: 90067

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,875

FILING DATE: July 12, 1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/102,385

FILING DATE: August 4, 1993
ATTORNEY/AGENT INFORMATION:

NAME: Oldenkamp, David J.
REGISTRATION NUMBER: 29,421

REFERENCE/DOCKET NUMBER: 111-197
TELECOMMUNICATION INFORMATION:

TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297

INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:

LENGTH: 340 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-446-875-16

Query Match 12.2%; Score 196.5; DB 2; Length 340;
Best Local Similarity 27.5%; Pred. No. 3.6e-13;
Matches 55; Conservative 32; Mismatches 60; Indels 51; Gaps 6;

QY 73 CRCAVSSSGOMLGSGLGAIDSAECVFRMNOAPTVGEADVQGRSTLRVSHSTVPL 132
DB 139 CRCAVNGSNGNRRESSYGEIDSHDFVLRMKAFTGFEADVGTCTHHLV----- 190
QY 133 LRNYSHYFOKARDTLVYMWGQGRHMD-----RVLGRTVYTL-----QLTRM 175
DB 191 ---YPSFRELGDVNSMILVFPKIDLEWVSAITTGITISHTYIPVPAKIRVQDKILY 247
QY 176 YPGLQVYTFERRMAYCDQIFQDETGKRRQSGSLSTGFWMTIMALECEIIVYGMS 235
DB 248 HPAFIKYVFNWLOQH-----GRYPTGILSVIFSMVACDEVLDYGFGA 291

QY 236 DSYCREKSHPSVPYHYFE 253
DB 292 DS--KGNMH-----HYME 302

RESULT 5
US-08-503-133A-2
Sequence 2, Application US/08503133A
Patent No. 5747326
GENERAL INFORMATION:
APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;
TITLE OF INVENTION: Isolated Polysialyl Transferases,
TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of
TITLE OF INVENTION: Production and Use
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felte & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/503,133A
FILING DATE: 17-JULY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/04289
FILING DATE: 22-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5747326man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: BOER 1050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-503-133A-2

Query Match 12.1%; Score 196.5; DB 1; Length 359;
Best Local Similarity 27.5%; Pred. No. 6.4e-13;
Matches 69; Conservative 39; Mismatches 98; Indels 45; Gaps 9;

QY 59 YSSVPQKPLVBRPCSCAVSSSGOMLGSGLGAIDSAECVFRMNOAPTVGEADVQGR 118
DB 125 HSLLPVSPMKRRRFTKCAVNGSGILDSGCKEIDSHNFVIRCNLAPVVEFAADVGT 164

QY 119 SLTRVSHSTVPL---LRNYS-----HYFOKARDTL-----YMWGQGRHMDRVYGR 164
DB 185 SDFITMNPVYQAFGFGFNRSDRAKVFHRLSMNLNSVLMIPAFMVGSGEKHVEW-NAL 243
QY 165 TYRLLIOLTMVYGLQV-----YTFERRMAYCDQIFQDETGKRRQSGSLSTGFWMT 218
DB 244 ILNKLKVRTAYPSLRLLIAVNRGIWLTNKVPI-----KRP-----STGLLMY 285
QY 219 ILALECEIIVYGMSDSYCREKSHPSVPYHYFEKGRLECOMYLAEQAAPSAHRFIT 278
DB 286 TLATRFDEIHLGFW--PPKDLNGKAVKHYHD---DLKRYTSNAPSHPMPLEFKT 339
QY 279 EKAVFSRMAK 289
DB 340 LNVILNRGALK 350

RESULT 6
US-08-576-775A-2
Sequence 2, Application US/08576775A
Patent No. 5849904
GENERAL INFORMATION:
APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;
TITLE OF INVENTION: Isolated Polysialyl Transferases,
TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of
TITLE OF INVENTION: Production and Use
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felte & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/576,775A
FILING DATE: 21-December-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/503,133
FILING DATE: 17-July-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/04289
FILING DATE: 22-December-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5849904man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: BOER 1050.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-576-775A-2

Query Match 12.1%; Score 196.5; DB 2; Length 359;
Best Local Similarity 27.5%; Pred. No. 6.4e-13;
Matches 69; Conservative 39; Mismatches 98; Indels 45; Gaps 9;

QY 59 YSSVPQKPLVBRPCSCAVSSSGOMLGSGLGAIDSAECVFRMNOAPTVGEADVQGR 118
DB 125 HSLLPVSPMKRRRFTKCAVNGSGILDSGCKEIDSHNFVIRCNLAPVVEFAADVGT 164

QY 119 STLRAVSHTSVPL--LNYS-----HYFKARDTL-----YVWGGRHMDRYLGR 164
 Db 185 SDFTMNPSPVQARFGGFNESPRAKVFHRLSLMNDLSVLMIPAFVWKGGEKHEWV-NAL 243
 QY 165 YRFTLLQLTRMYPGLQV-----YFTTERMAVCCQIFODETGKRRSGSFLSTGWFTM 218
 Db 244 ILKNKLKVTATVPSLRLIHAVRGYWLTKVPI-----KRP-----STGLMY 285
 QY 219 IIALCEIEIVVYGWSDSYCREKSHPSVYHYFEKGRLDECOMYLAEQAPRSARFTT 278
 Db 286 TLATRFCDLHLGFW--PPKDLNGKAVKHYTD---DLKRYFSNASPHRMPLEFXT 339
 QY 279 EKAVFSRWAK 289
 Db 340 LNVLHNRGALK 350

RESULT 7
 US-08-972-498-2
 ; Sequence 2, Application US/08972498
 ; Patent No. 5959078
 ; GENERAL INFORMATION:
 ; APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;
 ; APPLICANT: Nakayama, Jun; Eckhardt, Mathias
 ; TITLE OF INVENTION: Isolated Polysialyl Transferases,
 ; TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of
 ; TITLE OF INVENTION: Production and Use
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Felte & Lynch
 ; STREET: 805 Third Avenue
 ; CITY: New York City
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10022
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: PC-DOS
 ; SOFTWARE: Wordperfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/972,498
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/576,775
 ; FILING DATE: 21-December-1995
 ; APPLICATION NUMBER: 08/503,133
 ; FILING DATE: 17-July-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/EP94/04289
 ; FILING DATE: 22-December-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hanson, No. 5959078man D.
 ; REGISTRATION NUMBER: 30,946
 ; REFERENCE/DOCKET NUMBER: BOER 1050.1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 688-9200
 ; TELEFAX: (212) 838-3884
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 359 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-972-498-2

Query Match 12.1%; Score 196.5; DB 2; Length 359;
 Best Local Similarity 27.5%; Pred. No. 6.4e-13;
 Matches 69; Conservative 39; Mismatches 98; Indels 45; Gaps 9;
 QY 59 YSSVPDGKPLVREPCRCACAVVSSSGMGLGSLGAETDSACVFRMNQAPTVGFADVGQR 118

Db 125 HLLPEVSPMKRRRTKCAVGNISGLLDSGGCKEIDSNFTRCNLAPEVFEADVGTR 184
 QY 119 STLRAVSHTSVPL--LNYS-----HYFKARDTL-----YVWGGRHMDRYLGR 164
 Db 185 SDFTMNPSPVQARFGGFNESPRAKVFHRLSLMNDLSVLMIPAFVWKGGEKHEWV-NAL 243
 QY 165 YRFTLLQLTRMYPGLQV-----YFTTERMAVCCQIFODETGKRRSGSFLSTGWFTM 218
 Db 244 ILKNKLKVTATVPSLRLIHAVRGYWLTKVPI-----KRP-----STGLMY 285
 QY 219 IIALCEIEIVVYGWSDSYCREKSHPSVYHYFEKGRLDECOMYLAEQAPRSARFTT 278
 Db 286 TLATRFCDLHLGFW--PPKDLNGKAVKHYTD---DLKRYFSNASPHRMPLEFXT 339
 QY 279 EKAVFSRWAK 289
 Db 340 LNVLHNRGALK 350

RESULT 8
 US-08-899-545-2
 ; Sequence 2, Application US/08899545
 ; Patent No. 6020201
 ; GENERAL INFORMATION:
 ; APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;
 ; APPLICANT: Nakayama, Jun; Eckhardt, Mathias
 ; TITLE OF INVENTION: Isolated Polysialyl Transferases,
 ; TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of
 ; TITLE OF INVENTION: Production and Use
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Felte & Lynch
 ; STREET: 805 Third Avenue
 ; CITY: New York City
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10022
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: PC-DOS
 ; SOFTWARE: Wordperfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/899,545
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/503,133
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hanson, No. 6020201man D.
 ; REGISTRATION NUMBER: 30,946
 ; REFERENCE/DOCKET NUMBER: BOER 1050
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 688-9200
 ; TELEFAX: (212) 838-3884
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 359 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-899-545-2

Query Match 12.1%; Score 196.5; DB 3; Length 359;
 Best Local Similarity 27.5%; Pred. No. 6.4e-13;
 Matches 69; Conservative 39; Mismatches 98; Indels 45; Gaps 9;
 QY 59 YSSVPDGKPLVREPCRCACAVVSSSGMGLGSLGAETDSACVFRMNQAPTVGFADVGQR 118
 Db 125 HLLPEVSPMKRRRTKCAVGNISGLLDSGGCKEIDSNFTRCNLAPEVFEADVGTR 184

QY 119 STLRVSHTSVPL---LRRYS-----HYEOKARDTL-----YMWGGRHMDRVLCGR 164
DB 185 SDRITMPPSVQAFGGFNRSDRAKVRHLSMLNDVLMIPAPVWKGKHEWV-NAL 243
QY 165 YRTLLQLTRMYPGLQV-----YTFERRMAYCDQIFQDETKNRROSGSFLSTGFTM 218
DB 244 ILNKRLKVRAYSLRLIHAVRGYMLINKVPI-----KRP-----STGLIMY 285
QY 219 ILALICEELIVYGMVSDSYCREKSHPSVPHYFEKRLDECOMYLAHEQAPSAHFFIT 278
DB 286 TLATRCDEIHLXGFW--PPKDLNGKAVKHYDD---DLKRYFSPNAGSPHMPLEFKT 339
QY 279 EKAVFSRMAKX 289
DB 340 LNVILNRGALK 350

RESULT 9

US-08-446-875-2
Sequence 2, Application US/08446875
Patent No. 5858751
GENERAL INFORMATION:
APPLICANT: Paulson, James C.
APPLICANT: Wen, Xiaohong
APPLICANT: Livingston, Brian Duane
APPLICANT: Gillespie, William
APPLICANT: Kelm, Sorge
APPLICANT: Burlingame, Alma L.
APPLICANT: Medzhiradzky, Katalin
TITLE OF INVENTION: Compositions and Methods for the
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pome, Smith, Lande & Rose
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,875
FILING DATE: July 12, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/102,385
FILING DATE: August 4, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Oldenkamp, David J.
REGISTRATION NUMBER: 29,421
TELEPHONE: (310) 788-5000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 277-1297
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 343 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-446-875-2

Query Match 11.7%; Score 191; DB 2; Length 343;
Best Local Similarity 30.0%; Pred. No. 2,4e-12;
Matches 57; Conservative 36; Mismatches 77; Indels 20; Gaps 5;
QY 67 PLVRE---PGRCAVSSSGOMLGSGLGAIDSAECVFRMNOAPTYGFADVGQRSTLRV 123
DB 133 PLEKRLVSCRCRCAVVGNSGNLKESTYYGPQIDSHDFVLRNKAFTBGFADVGSKTTHHF 192

QY 124 VSHTSVPLLRNYSHFQKARDTLVWVGGRHMDRVLCGRTRYRTLLQLTRMYPGLQVYT 183
DB 193 VYPSFRELAEQVS-----MILVPEFTDLEWISATTTGTTSHTYVPVPA-KIKV 242
QY 184 FTERMAYCDQIFQDETKNRROSGSFLSTGFTMLALICEELIVYGMVSDSYCREKS 243
DB 243 KKEKILYHPAFIKVYFDRMLQGHGRVPSGTILSVFSLHICDEVDLGRGADS--KGNW 300
QY 244 HPSVPHYFE 253
DB 301 H-----HYWE 305

RESULT 10

US-08-102-385G-2
Sequence 2, Application US/08102385G
Patent No. 5962294
GENERAL INFORMATION:
APPLICANT: Paulson, James C.
APPLICANT: Wen, Xiaohong
APPLICANT: Livingston, Brian Duane
APPLICANT: Gillespie, William
APPLICANT: Kelm, Sorge
APPLICANT: Burlingame, Alma L.
APPLICANT: Medzhiradzky, Katalin
TITLE OF INVENTION: Compositions and Methods for the
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
STREET: 2029 Century Park East, 38th Floor
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/102,385G
FILING DATE: 04-AUG-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/925369
FILING DATE: 04-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Oldenkamp, David J.
REGISTRATION NUMBER: 29,421
TELEPHONE: (310) 788-5000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 277-1297
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 343 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-102-385G-2

Query Match 11.7%; Score 191; DB 2; Length 343;
Best Local Similarity 30.0%; Pred. No. 2,4e-12;
Matches 57; Conservative 36; Mismatches 77; Indels 20; Gaps 5;
QY 67 PLVRE---PGRCAVSSSGOMLGSGLGAIDSAECVFRMNOAPTYGFADVGQRSTLRV 123
DB 133 PLEKRLVSCRCRCAVVGNSGNLKESTYYGPQIDSHDFVLRNKAFTBGFADVGSKTTHHF 192
QY 124 VSHTSVPLLRNYSHFQKARDTLVWVGGRHMDRVLCGRTRYRTLLQLTRMYPGLQVYT 183

Db 193 VYSESFRELAQEV-----MILVFKTDLBENVIASATTTGTSHTYVPA-KIKV 242

QY 184 FTFRMAYCQITQDETGKRRSGSLSTGWTMLALCEIIVYGWVSYSYCKRS 243

Db 243 KKEKILYHFAFKYVDFRWLOCHGRYPSTGILSVIFSLHCDEVDLYFGADS--KGNW 300

QY 244 HPSVPYHYFE 253

Db 301 H-----HYWE 305

RESULT 11

US-08-626-994A-3

; Sequence 3, Application US/08626994A

; Patent No. 5798244

; GENERAL INFORMATION:

; APPLICANT: Shuichi TSUJI et al.

; TITLE OF INVENTION: Sia' 2,3gala 1,4GlcNAc ' 2,8-

; TITLE OF INVENTION: SIATYLTANSFERASE

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Greenblum & Bernstein, P.L.C.

; STREET: 1941 Roland Clarke Place

; City: Reston

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 20191

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 1.44 diskette

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: Word Perfect 5.1+ (ASCII)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/626,994A

; FILING DATE: April 3, 1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 77469/1995

; FILING DATE: April 3, 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Arnold Turk

; REGISTRATION NUMBER: 33,094

; REFERENCE/DOCKET NUMBER: P14595

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 716-1191

; TELEFAX: (703) 716-1180

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 339 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: linear

US-08-626-994A-3

Query Match 11.7%; Score 190.5; DB 1; Length 339;

Best Local Similarity 30.0%; Pred. No. 2,7e-12;

Matches 65; Conservative 26; Mismatches 81; Indels 43; Gaps 8;

QY 60 SSVPDGRPLVREPCRCGAVVSSGOMLGSLGAEIDSACEVFRMNQAPTYGFADVGORS 119

Db 105 SLIPDVSPIMKRRYVCAVVGNSGILTGSCQGEIDKSDVSRCNFAPTEAFHKDVGRKT 164

QY 120 TLKRVSHTSVPLLRNYSHYFQKARDTL-----YMWGQGRHMDRVLGRTYRT 168

Db 165 NLT-----TNPSTLEKYYNNLLTIQDRNNFFLSKLKLDGAILW-----IPAFFHT 211

QY 169 LLQLTRMYPGLQYTYFTER-----MMAYCQIFQ--DETGKRRSGSLSTGWTML 220

Db 212 SATVTRTL-----VDFVEHGRQLKVQLAMPGNIMQHVNRWKAKHLSPKRLSTGILMYTL 267

QY 221 ALCEIIVVYGWVS--DSYCKRKHSPVYHYFEK 254

Db 268 ASAICEIHLXGFWPFGFDPNTR-----DLPYHYDK 300

RESULT 12

US-08-957-742-3

; Sequence 3, Application US/08957742

; Patent No. 6017743

; GENERAL INFORMATION:

; APPLICANT: Shuichi TSUJI et al.

; TITLE OF INVENTION: Sia' 2,3gala 1,4GlcNAc ' 2,8-

; TITLE OF INVENTION: SIATYLTANSFERASE

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Greenblum & Bernstein, P.L.C.

; STREET: 1941 Roland Clarke Place

; City: Reston

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 20191

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 1.44 diskette

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: Word Perfect 5.1+ (ASCII)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/957,742

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/626,994

; FILING DATE: April 3, 1996

; APPLICATION NUMBER: JP 77469/1995

; FILING DATE: April 3, 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Arnold Turk

; REGISTRATION NUMBER: 33,094

; REFERENCE/DOCKET NUMBER: P14595

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 716-1191

; TELEFAX: (703) 716-1180

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 339 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: linear

US-08-957-742-3

Query Match 11.7%; Score 190.5; DB 3; Length 339;

Best Local Similarity 30.0%; Pred. No. 2,7e-12;

Matches 65; Conservative 28; Mismatches 81; Indels 43; Gaps 8;

QY 60 SSVPDGRPLVREPCRCGAVVSSGOMLGSLGAEIDSACEVFRMNQAPTYGFADVGORS 119

Db 105 SLIPDVSPIMKRRYVCAVVGNSGILTGSCQGEIDKSDVSRCNFAPTEAFHKDVGRKT 164

QY 120 TLKRVSHTSVPLLRNYSHYFQKARDTL-----YMWGQGRHMDRVLGRTYRT 168

Db 165 NLT-----TNPSTLEKYYNNLLTIQDRNNFFLSKLKLDGAILW-----IPAFFHT 211

QY 169 LLQLTRMYPGLQYTYFTER-----MMAYCQIFQ--DETGKRRSGSLSTGWTML 220

Db 212 SATVTRTL-----VDFVEHGRQLKVQLAMPGNIMQHVNRWKAKHLSPKRLSTGILMYTL 267

QY 221 ALCEIIVVYGWVS--DSYCKRKHSPVYHYFEK 254

Db 268 ASAICEIHLXGFWPFGFDPNTR-----DLPYHYDK 300

RESULT 13

US-08-626-994A-1

; Sequence 1, Application US/08626994A

; Patent No. 5798244

; GENERAL INFORMATION:

```

; APPLICANT: Shuichi TSUJI et al.
; TITLE OF INVENTION: Sia' 2,3gala 1,4GlcNAc ` 2,8-
; TITLE OF INVENTION: SIALYLTRANSFERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenblum & Bernstein, P.L.C.
; STREET: 1941 Roland Clarke Place
; CITY: Reston
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 20191
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1+ (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/626,994A
; FILING DATE: April 3, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 77469/1995
; FILING DATE: April 3, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold Turk
; REGISTRATION NUMBER: 33,094
; REFERENCE/DOCKET NUMBER: P14595
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 716-1191
; TELEFAX: (703) 716-1180
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
;
; US-08-626-994A-1
;
; Query Match 11.7%; Score 190.5; DB 1; Length 364;
; Best Local Similarity 30.0%; Pred. No. 3e-12;
; Matches 65; Conservative 28; Mismatches 81; Indels 43; Gaps 8;
;
; QY 60 SSVPDCKELVREPCRCGAVVSSSGOMLGSGLGAEIDSAECVFRNNOAPTYGFADVQGRS 119
; DB 130 SLIPDVSPINMKKRYNVCVAVNGSILTGSGCGEIDKSDVSRNCFAPTEAFHDKVGRKT 189
; QY 120 TLRVSHTSVPLLRNYSHYFOKARDTL-----YMWGQGRHMDRVLGRTYRT 168
; DB 190 NLT-----TNPSTLEKLYNNLTIQDNFFLSIKLDGAILW-----IPAFFHT 236
; QY 169 LLOLTRNYPGLQVYTFER-----MMAYCDQIFQ--DETGKRRSGSFLSTGWFMTIL 220
; DB 237 SATVTRTL-----VDFEYHGRQLKVLQWLPNGNIMQHVNRYWKXKHLSPKSLSTGILMYTL 292
;
; QY 221 ALBLCEIIVYGVNS---DSYCREKSHPSVPHYHFEK 254
; DB 293 ASAIICEIHLYGFWPGFDPNTRB---DLPYHYDK 325
;
; RESULT 14
; US-08-957-742-1
; Sequence 1, Application US/08957742
; Patent No. 6017743
; GENERAL INFORMATION:
; APPLICANT: Shuichi TSUJI et al.
; TITLE OF INVENTION: Sia' 2,3gala 1,4GlcNAc ` 2,8-
; TITLE OF INVENTION: SIALYLTRANSFERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenblum & Bernstein, P.L.C.
; STREET: 1941 Roland Clarke Place
; CITY: Reston
; STATE: Virginia

```

```

; COUNTRY: U.S.A.
; ZIP: 20191
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1+ (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/957,742
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/626,994
; FILING DATE: April 3, 1996
; APPLICATION NUMBER: JP 77469/1995
; FILING DATE: April 3, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold Turk
; REGISTRATION NUMBER: 33,094
; REFERENCE/DOCKET NUMBER: P14595
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 716-1191
; TELEFAX: (703) 716-1180
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
;
; US-08-957-742-1
;
; Query Match 11.7%; Score 190.5; DB 3; Length 364;
; Best Local Similarity 30.0%; Pred. No. 3e-12;
; Matches 65; Conservative 28; Mismatches 81; Indels 43; Gaps 8;
;
; QY 60 SSVPDCKELVREPCRCGAVVSSSGOMLGSGLGAEIDSAECVFRNNOAPTYGFADVQGRS 119
; DB 130 SLIPDVSPINMKKRYNVCVAVNGSILTGSGCGEIDKSDVSRNCFAPTEAFHDKVGRKT 189
; QY 120 TLRVSHTSVPLLRNYSHYFOKARDTL-----YMWGQGRHMDRVLGRTYRT 168
; DB 190 NLT-----TNPSTLEKLYNNLTIQDNFFLSIKLDGAILW-----IPAFFHT 236
; QY 169 LLOLTRNYPGLQVYTFER-----MMAYCDQIFQ--DETGKRRSGSFLSTGWFMTIL 220
; DB 237 SATVTRTL-----VDFEYHGRQLKVLQWLPNGNIMQHVNRYWKXKHLSPKSLSTGILMYTL 292
;
; QY 221 ALBLCEIIVYGVNS---DSYCREKSHPSVPHYHFEK 254
; DB 293 ASAIICEIHLYGFWPGFDPNTRB---DLPYHYDK 325
;
; RESULT 15
; US-08-666-367B-6
; Sequence 6, Application US/08666367B
; Patent No. 5854042
; GENERAL INFORMATION:
; APPLICANT: Shuichi TSUJI et al.
; TITLE OF INVENTION: NOVEL SUGAR-CHAIN SYNTHETASE AND PROCESS FOR
; TITLE OF INVENTION: PRODUCING THE SAME
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoch, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/666,367B
FILING DATE: August 19, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-666-367B-6

Query Match 10.9%; Score 178; DB 2; Length 355;
Best Local Similarity 25.2%; Pred. No. 6.8e-11;
Matches 63; Conservative 37; Mismatches 80; Indels 70; Gaps 9;
QY 73 CRCAVYSSSGQMIGSLGAEIDSAECVFRMNQAPTVGFADYVGRSTLRVSHSVPLL 132
DB 126 CISCAYVNGGIGLNNSGMGEIDSHDYFRVSGAVIKGIEKDVGTSTFGFTAYSLVSS 185
QY 133 LRNSHY-FOKAPDLYVMWGQGRHMDRVLGRTYRTLLQITRNYPGIQQV----- 181
DB 186 LQNGHKGFKI-----PQKHI-----RYIHLEAVRDYEWIKALLDDKDIRKG 230
QY 182 -----YFTFRMMAYCDQIFQDETGNRRQSSFL-----STGW 215
DB 231 FLNYGRRPRERFDEFTMNRYLVAHDPFL--RYLKNRFLKSKOLQKPYRLRYPTTGA 287
QY 216 FTMLALELCEIIVYGVSDSYCREKSHPSVPYHYFEKGRLDCEOMYLAHEQAPRSAMR 275
DB 288 LLLLTALHLCDRVSAKGYITE-----GHQKYSDDHYDK-EMKRLVVFVYVNH----- 332
QY 276 FITEKAVPSR 285
DB 333 FNLEKQVWKR 342

Search completed: December 21, 2004, 22:44:59
Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2004, 22:13:05 ; Search time 77 Seconds

(without alignments)
1403.457 Million cell updates/sec

Title: US-10-019-114A-7

Perfect score: 1626
Sequence: 1 MKRPGRLVLIILCSVPSAV.....FSRMKKRPVFAHPKRWTE 302Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1589859 seqs, 357834939 residues

Total number of hits satisfying chosen parameters: 1589859

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA:*

1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1626	100.0	302	11	US-09-978-360A-407
2	1626	100.0	302	14	US-10-315-664-75
3	1622	99.8	302	10	US-09-895-298-66
4	1622	99.8	302	10	US-09-895-298-120
5	1167	71.8	218	15	US-10-262-511-184
6	726.5	44.7	305	17	US-10-844-874-20
7	716.5	44.1	305	11	US-09-833-245-1914
8	716.5	44.1	305	11	US-09-833-245-1915
9	714.5	43.9	305	15	US-10-188-186-22
10	713.5	43.9	305	14	US-10-104-047-2350
11	694	42.7	318	15	US-10-188-186-24
12	677.5	41.7	305	15	US-10-188-186-20
13	639.5	39.3	336	14	US-10-097-340-189

14	639.5	39.3	336	14	US-10-094-749-2441	Sequence 2441, Ap
15	636.5	39.1	336	14	US-10-288-252-11	Sequence 11, Appl
16	636.5	39.1	336	10	US-10-398-038-11	Sequence 11, Appl
17	537	33.0	299	10	US-09-946-374-56	Sequence 56, Appl
18	537	33.0	299	13	US-10-052-586-206	Sequence 206, App
19	537	33.0	299	14	US-10-174-590-206	Sequence 206, App
20	537	33.0	299	14	US-10-176-758-206	Sequence 206, App
21	537	33.0	299	14	US-10-175-737-206	Sequence 206, App
22	537	33.0	299	14	US-10-174-581-206	Sequence 206, App
23	537	33.0	299	14	US-10-176-483-206	Sequence 206, App
24	537	33.0	299	14	US-10-176-749-206	Sequence 206, App
25	537	33.0	299	14	US-10-176-914-206	Sequence 206, App
26	537	33.0	299	14	US-10-176-915-206	Sequence 206, App
27	537	33.0	299	14	US-10-173-706-206	Sequence 206, App
28	537	33.0	299	14	US-10-175-738-206	Sequence 206, App
29	537	33.0	299	14	US-10-175-752-206	Sequence 206, App
30	537	33.0	299	14	US-10-176-482-206	Sequence 206, App
31	537	33.0	299	14	US-10-176-757-206	Sequence 206, App
32	537	33.0	299	14	US-10-176-913-206	Sequence 206, App
33	537	33.0	299	14	US-10-180-552-206	Sequence 206, App
34	537	33.0	299	14	US-10-180-557-206	Sequence 206, App
35	537	33.0	299	14	US-10-173-700-206	Sequence 206, App
36	537	33.0	299	14	US-10-174-572-206	Sequence 206, App
37	537	33.0	299	14	US-10-174-579-206	Sequence 206, App
38	537	33.0	299	14	US-10-174-582-206	Sequence 206, App
39	537	33.0	299	14	US-10-174-588-206	Sequence 206, App
40	537	33.0	299	14	US-10-175-739-206	Sequence 206, App
41	537	33.0	299	14	US-10-175-740-206	Sequence 206, App
42	537	33.0	299	14	US-10-175-743-206	Sequence 206, App
43	537	33.0	299	14	US-10-176-488-206	Sequence 206, App
44	537	33.0	299	14	US-10-176-492-206	Sequence 206, App
45	537	33.0	299	14	US-10-176-747-206	Sequence 206, App

ALIGNMENTS

RESULT 1
US-09-978-360A-407 Application US/09978360A
; Sequence 407, Application US/09978360A
; Publication No. US20040110939A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymrie
; APPLICANT: Bouguetieret, Lydie
; APPLICANT: Joberet, Severin
; APPLICANT: Clusel, Catherine
; TITLE OR INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56.USA.CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810

SOFTWARE: Patent.pm
 SEQ ID NO 407
 LENGTH: 302
 TYPE: PRT
 ORGANISM: Homo Sapiens
 FEATURE:
 NAME/KEY: SIGNAL
 LOCATION: -18...-1
 US-09-978-360A-407

Query Match 100.0%; Score 1626; DB 11; Length 302;
 Best Local Similarity 100.0%; Pred. No. 2.4e-160;
 Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAPGRVLVILICSVFSAVYIILCCWAGLPICLATCLDHHFPTGSRPTVGPPLHFSGY 60
 DB 1 MKAPGRVLVILICSVFSAVYIILCCWAGLPICLATCLDHHFPTGSRPTVGPPLHFSGY 60
 QY 61 SVPDGKPLVREPCRCSCAVSSSGQMLGSGLAETIDSAECVFRMNOAPTGVFEADVQGRST 120
 DB 61 SVPDGKPLVREPCRCSCAVSSSGQMLGSGLAETIDSAECVFRMNOAPTGVFEADVQGRST 120
 QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYMWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
 DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYMWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
 QY 181 VYTFTERMMAVCDQIFODETGKRRQSGSLSTGMFTMLALELCEIIVYGVMSDSYCR 240
 DB 181 VYTFTERMMAVCDQIFODETGKRRQSGSLSTGMFTMLALELCEIIVYGVMSDSYCR 240
 QY 241 EKSHPSVPYHYFEKGRLEDCOMYLAEQAPRSARHPTTEKAVFSRWAKKRPVFAHPSWR 300
 DB 241 EKSHPSVPYHYFEKGRLEDCOMYLAEQAPRSARHPTTEKAVFSRWAKKRPVFAHPSWR 300
 QY 301 TE 302
 DB 301 TE 302

RESULT 2

US-10-315-664-75
 Sequence 75, Application US/10315664
 Publication No. US20030203377A1
 GENERAL INFORMATION:
 APPLICANT: Dumas Milne Edwards, J.B.
 APPLICANT: Bougueleret, L.
 APPLICANT: Jobert, S.
 TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal
 FILE REFERENCE: GENSET.050CP3
 CURRENT APPLICATION NUMBER: US/10/315,664
 PRIOR FILING DATE: 2002-12-09
 PRIOR APPLICATION NUMBER: US/09/599,360
 PRIOR FILING DATE: 2000-06-21
 PRIOR APPLICATION NUMBER: 60/113,686
 PRIOR FILING DATE: 1998-12-22
 PRIOR APPLICATION NUMBER: 60/141,032
 PRIOR FILING DATE: 1999-06-25
 PRIOR APPLICATION NUMBER: 09/469,099
 PRIOR FILING DATE: 1999-12-21
 NUMBER OF SEQ ID NOS: 123
 SOFTWARE: Patent.pm
 SEQ ID NO 75
 LENGTH: 302
 TYPE: PRT
 ORGANISM: Homo Sapiens
 FEATURE:
 NAME/KEY: SIGNAL
 LOCATION: -18...-1
 US-10-315-664-75

Query Match 100.0%; Score 1626; DB 14; Length 302;
 Best Local Similarity 100.0%; Pred. No. 2.4e-160;

Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAPGRVLVILICSVFSAVYIILCCWAGLPICLATCLDHHFPTGSRPTVGPPLHFSGY 60
 DB 1 MKAPGRVLVILICSVFSAVYIILCCWAGLPICLATCLDHHFPTGSRPTVGPPLHFSGY 60
 QY 61 SVPDGKPLVREPCRCSCAVSSSGQMLGSGLAETIDSAECVFRMNOAPTGVFEADVQGRST 120
 DB 61 SVPDGKPLVREPCRCSCAVSSSGQMLGSGLAETIDSAECVFRMNOAPTGVFEADVQGRST 120
 QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYMWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
 DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYMWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
 QY 181 VYTFTERMMAVCDQIFODETGKRRQSGSLSTGMFTMLALELCEIIVYGVMSDSYCR 240
 DB 181 VYTFTERMMAVCDQIFODETGKRRQSGSLSTGMFTMLALELCEIIVYGVMSDSYCR 240
 QY 241 EKSHPSVPYHYFEKGRLEDCOMYLAEQAPRSARHPTTEKAVFSRWAKKRPVFAHPSWR 300
 DB 241 EKSHPSVPYHYFEKGRLEDCOMYLAEQAPRSARHPTTEKAVFSRWAKKRPVFAHPSWR 300
 QY 301 TE 302
 DB 301 TE 302

RESULT 3

US-09-895-298-66
 Sequence 66, Application US/09895298
 Publication No. US20030078405A1
 GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: 47 Human Secreted Proteins
 FILE REFERENCE: P2035P1
 CURRENT APPLICATION NUMBER: US/09/895,298
 PRIOR FILING DATE: 2001-07-02
 PRIOR APPLICATION NUMBER: 09/591,16
 PRIOR FILING DATE: 2000-06-09
 PRIOR APPLICATION NUMBER: PCT/US99/29950
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: 60/113,006
 PRIOR FILING DATE: 1998-12-18
 PRIOR APPLICATION NUMBER: 60/112,809
 PRIOR FILING DATE: 1998-12-17
 NUMBER OF SEQ ID NOS: 231
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 66
 LENGTH: 302
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (237)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-09-895-298-66

Query Match 99.8%; Score 1622; DB 10; Length 302;
 Best Local Similarity 99.7%; Pred. No. 6.2e-160;
 Matches 301; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKAPGRVLVILICSVFSAVYIILCCWAGLPICLATCLDHHFPTGSRPTVGPPLHFSGY 60
 DB 1 MKAPGRVLVILICSVFSAVYIILCCWAGLPICLATCLDHHFPTGSRPTVGPPLHFSGY 60
 QY 61 SVPDGKPLVREPCRCSCAVSSSGQMLGSGLAETIDSAECVFRMNOAPTGVFEADVQGRST 120
 DB 61 SVPDGKPLVREPCRCSCAVSSSGQMLGSGLAETIDSAECVFRMNOAPTGVFEADVQGRST 120
 QY 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYMWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180
 DB 121 LRVVSHTSVPLLRNYSHYFOKARDTLVYMWGGRHMDRVLGRTYRTLLQLTRMYPGLQ 180

QY 181 VYFTEHMAVCDQIFODETGKRRSGSFLSTGCMFTMLALECEETIVYGVMSDYCR 240
DB 181 VYFTEHMAVCDQIFODETGKRRSGSFLSTGCMFTMLALECEETIVYGVMSDYCR 240
QY 241 EKSHPSVPHYFEKGRUDECOMYLAEQAPRSARHFTTEKAVFSRMKKRPVFAHPSMR 300
DB 241 EKSHPSVPHYFEKGRUDECOMYLAEQAPRSARHFTTEKAVFSRMKKRPVFAHPSMR 300
QY 301 TE 302
DB 301 TE 302

RESULT 4

US-09-895-298-120
Sequence 120, Application US/09895298
Publication No. US20030078405A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 47 Human Secreted Proteins
FILE REFERENCE: P2035P1
CURRENT APPLICATION NUMBER: US/09/895,298
CURRENT FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: 09/591,16
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: PCT/US99/29950
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: 60/113,006
PRIOR FILING DATE: 1998-12-18
PRIOR APPLICATION NUMBER: 60/112,809
PRIOR FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 231
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 120
LENGTH: 345
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (280)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-895-298-120

Query Match 99.8%; Score 1622; DB 10; Length 345;
Best Local Similarity 99.7%; Pred. No. 7.5e-160;
Matches 301; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKAPGRVLIIILGSVPSAVYIILCCWAGPLICLATCLDHHFPTGSRPTVGPPLHPSGYS 60
DB 44 MKAPGRVLIIILGSVPSAVYIILCCWAGPLICLATCLDHHFPTGSRPTVGPPLHPSGYS 103
QY 61 SVDDGKPLVNEPCRSQAVSSSGMLGSGGAEIDSACVFRMNAQATVGFADVGQRST 120
DB 104 SVDDGKPLVNEPCRSQAVSSSGMLGSGGAEIDSACVFRMNAQATVGFADVGQRST 163
QY 121 LRVVSHTSVPLLRNLSHYFOKARDTLVYMWGQGRHMDRVLGRTYFTLLQLTRMYPGLQ 180
DB 164 LRVVSHTSVPLLRNLSHYFOKARDTLVYMWGQGRHMDRVLGRTYFTLLQLTRMYPGLQ 223
QY 181 VYFTEHMAVCDQIFODETGKRRSGSFLSTGCMFTMLALECEETIVYGVMSDYCR 240
DB 224 VYFTEHMAVCDQIFODETGKRRSGSFLSTGCMFTMLALECEETIVYGVMSDYCR 283
QY 241 EKSHPSVPHYFEKGRUDECOMYLAEQAPRSARHFTTEKAVFSRMKKRPVFAHPSMR 300
DB 284 EKSHPSVPHYFEKGRUDECOMYLAEQAPRSARHFTTEKAVFSRMKKRPVFAHPSMR 343
QY 301 TE 302
DB 344 TE 345

RESULT 5

US-10-262-511-184
Sequence 184, Application US/10262511
Publication No. US20040038223A1
GENERAL INFORMATION:
APPLICANT: Smithson, Glenna
APPLICANT: Millet, Isabelle
APPLICANT: Peyman, John A.
APPLICANT: Kekuda, Ramesh
APPLICANT: Ju, Jingfang
APPLICANT: Li, Li
APPLICANT: Guo, Xiaojia (Sasha)
APPLICANT: Patturajan, Meera
APPLICANT: Spytek, Kimberly A.
APPLICANT: Edinger, Shlomit R.
APPLICANT: Ellerman, Karen
APPLICANT: Malyankar, Uriel M.
APPLICANT: Ort, Tatiana
APPLICANT: Gorman, Linda
APPLICANT: Zernusen, Bryan D.
APPLICANT: Anderson, David W.
APPLICANT: Zhong, Mei
APPLICANT: Caterton, Elina
APPLICANT: Miller, Charles E.
APPLICANT: Rastelli, Luca
APPLICANT: Stone, David J.
APPLICANT: Pena, Carol E. A.
APPLICANT: Shenoy, Suresh G.
APPLICANT: Shinkets, Richard A.
APPLICANT: Rothenberg, Mark E.
APPLICANT: Leach, Martin D.
APPLICANT: Agee, Michele L.
APPLICANT: Berghe, Constance

TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

FILE REFERENCE: 21402-462C

CURRENT APPLICATION NUMBER: US/10/262,511

PRIOR FILING DATE: 2003-05-28

PRIOR APPLICATION NUMBER: 60/326,483

PRIOR FILING DATE: 2001-10-02

PRIOR APPLICATION NUMBER: 60/373,815

PRIOR FILING DATE: 2002-04-19

PRIOR APPLICATION NUMBER: 60/327,917

PRIOR FILING DATE: 2001-10-09

PRIOR APPLICATION NUMBER: 60/381,642

PRIOR FILING DATE: 2002-05-17

PRIOR APPLICATION NUMBER: 60/328,029

PRIOR FILING DATE: 2002-10-09

PRIOR APPLICATION NUMBER: 60/381,038

PRIOR FILING DATE: 2002-05-16

PRIOR APPLICATION NUMBER: 60/328,056

PRIOR FILING DATE: 2001-10-09

PRIOR APPLICATION NUMBER: 60/373,260

PRIOR FILING DATE: 2002-04-17

PRIOR APPLICATION NUMBER: 60/373,826

PRIOR FILING DATE: 2002-04-19

PRIOR APPLICATION NUMBER: 60/327,435

PRIOR FILING DATE: 2001-10-05

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 439

SOFTWARE: Curaseqdist version 0.1

SEQ ID NO 184

LENGTH: 218

TYPE: PRT

ORGANISM: Homo sapiens

US-10-262-511-184

Query Match 71.8%; Score 1167; DB 15; Length 218;
Best Local Similarity 100.0%; Pred. No. 9.2e-113;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 MLGSGGAEIDSACVFRMNAQATVGFADVGQRSTLRVVSHTSVPLLRNLSHYFOKAR 144
DB 1 MLGSGGAEIDSACVFRMNAQATVGFADVGQRSTLRVVSHTSVPLLRNLSHYFOKAR 60

Qy	175	MPGLOVYVTFETRMMAVYCOQIFODEGKGRKRGSGSFLSTGCMFMIATLETCEEIVVYGVN	234
Db	179	IYPNQIVYVTTTEKMSYCGDGVFKKETGRDVRVGGSLSTGCMFLFILMDACVGIHYVGM	238
Qy	235	SDSYCREKSHSPVRYHYHFEKRGILDECOMVLAHQAPRSARHFTTEKAVPSRMKKRPV	294
Db	239	NDTYCKTEGVRKVPYHYEGR-DECDSYFLHENAIVGGHRTTEKAVPEAKMKRRILP	297
Qy	295	AHPSTW 299	
Db	298	THPNW 302	

```

RESULT 9
US-10-188-186-22
; Sequence 22, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 22
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-188-186-22

```

	Query Match	43.9%	Score 714.5,	DB 15,	Length 305,
	Best Local Similarity	52.2%	Pred. No.1.9e-65,		
	Matches 128;	Conservative 49;	Mismatches 65;	Indels 3;	Gaps 3;
QY	55 HFSGSYSSVDEKPKPLVRBPCRSACAVVSSSGOMLGSLGAGIEIDSABECVFPMNQAPTVGSPAD	114			
Db	61 HY-GYINVKTOEPFLD-CDLCAIVSNGQWGVQKVEIDRSSCTIMENNAAPTGYEED	118			
QY	115 VGORSTLRVHTSVPLLRLRYSHFQCARPTLVYWQQGHMDRVLGGRTYRTLLOLTR	174			
Db	119 VGRMTLRVHSTSVPLLLKNPDYLFKKANLTIIYIWDPFRNRKDGNIGVIYNMLAKTVG	178			
QY	175 MYPGLOVYTFTERMMAYCDQIFODETGKRBRQSGSFILSTGMFTMIALLECEIIVYGAV	234			
Db	179 IYPNAQIVITTEKMSYCDGVFKKETGKDRVQSGSYLSTGTFTFLAMDACGIHYGMI	238			
QY	235 SDSVCREKSHSPVPHFYHEKRLDECMNYLAHQAPRAHRPITEKAVFSMAKRRPIVF	294			
Db	239 NDTICKRGKYAKVPVHYEQGR-DECDERYFLHEAHAPYGCHRFITEKVFAAKKHRIIF	297			
QY	295 AHPSM 299				

Db 298 THPNW 302

```

RESULT 10
US-10-104-047-2350
; Sequence 2350, Application US//10104047
; Publication NO. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIORITY APPLICATION NUMBER:
; PRIORITY FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2350
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-104-047-2350

```

	Query Match	43.9%	Score 713.5	DB 14	Length 305
	Best Local Similarity	52.2%	Pred. No. 2,46	DB 15	
	Matches 128	Conservative 49	Mismatches 65	Indels 3	Gaps 3
Qy	55 HFSGTSVSDPGKPLVREPFRCSCAVVSSSGMGLGALNEIDSAECVFRNAPLTVGFEAD	114			
Db	61 HY-GYINWVKQEPFLQD-CDLCALVNSGQWGVQKVENEDRSSCIWRMNNAPIKGVED	118			
Qy	115 VGORSTLRVVSHTSVPLLRNYSHPKARDTLVWVGGRHMRVLGGRYRTLLQTLR	174			
Db	119 VGRMTLMRVVSHSVPLLRNKPDPVFEKAEATTIIVIMGPRFNMRKDGNGIILYNNMLKTVG	178			
Qy	175 MYRGLQVYTTTERMMAYCDQIFQDEGTGNRRQSSFSFTGFTMTALAECEELVYGVV	234			
Db	179 IYPAQAQIVYVTTTEKMSYCDGVFKKETKDRQSSYSTGFTFLAMDACYGIHYGMI	238			
Qy	235 SDSYCREKSHPSVYHYFEKGRGLDECOMYLAEQAPRSARHFRITFEKAFVSMAKRRPIVF	294			
Db	239 NDLYCKNEGRKXVYHYHEQGR-DECEYFLHEHAPVGGHRFRITFEKVFPAWMAKGRHIF	297			
Qy	295 AHPBW 299				
Db	298 THPW 302				

```

RESULT 11
US-10-188-24
: Sequence 24, Application US/10188186
: Publication No. US20040029789A1
: GENERAL INFORMATION:
: APPLICANT: Anderson et al.
: TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: 21402-397C
: CURRENT APPLICATION NUMBER: US/10/188,186
: CURRENT FILING DATE: 2002-07-02
: PRIOR APPLICATION NUMBER: 60/303046
: PRIOR FILING DATE: 2001-07-05
: PRIOR APPLICATION NUMBER: 60/360814
: PRIOR FILING DATE: 2002-03-01
: PRIOR APPLICATION NUMBER: 60/303828
: PRIOR FILING DATE: 2001-09-07
: PRIOR APPLICATION NUMBER: 60/323380
: PRIOR FILING DATE: 2001-09-19
: PRIOR APPLICATION NUMBER: 60/361133
: PRIOR FILING DATE: 2002-03-01
: PRIOR APPLICATION NUMBER: 60/304016
: PRIOR FILING DATE: 2001-07-09
: PRIOR APPLICATION NUMBER: 60/304502
: PRIOR FILING DATE: 2001-07-11

```

```
;; PRIOR APPLICATION NUMBER: 60/305262
;; PRIOR FILING DATE: 2001-07-13
;; PRIOR APPLICATION NUMBER: 60/373881
;; PRIOR FILING DATE: 2002-04-19
;; PRIOR APPLICATION NUMBER: 60/305673
;; PRIOR FILING DATE: 2001-07-16
;; Remaining Prior Application data removed - See File Wrapper or PAM.
;; NUMBER OF SEQ ID NOS: 368
;; SOFTWARE: Custom
US-10-188-186-24

Query Match      42.7%; Score 694; DB 15; Length 318;
Best Local Similarity 49.6%; Pred. No. 2.8e-63; Indels 16; Gaps 4;
Matches 128; Conservative 48; Mismatches 66;

QY 55 HFGSYSSVDDGKPLVREPCRSQAVSSSGQMLGSLGAEIDSACVFRMNAQPTVGFAD 114
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 61 HY-GYINVKTOEPLQD-CDLCAIVNSGQWGVKGNEDRSSCIWRMNNAPTGYEED 118
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 115 VGORSTLRVSHTSVPLLRNYSHPQAKADTLVMWQGRHMDVLGRTYRTLLQTR 174
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 119 VGRMTMIRVSHTSVPLLRNYSHPQAKADTLVMWQGRHMDVLGRTYRTLLQTR 178
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 175 MYPGLOVTFEPRMMAVCDQIFQDETGNR-----RQSGSFLSTGWTMLA 221
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 179 IYPNAQIVTTEKMSYCDGVFKETGDSVLTLSQSSPOLYQSSSYLSTGFTILA 238
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 222 LELCEIIVVGMVSDSYCREKSHSPVPHYFEKRLDECQMYLAHEQAPRSARHFTIEKA 281
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 239 MDACVGHVGMINDYTKTEGYRKVPYHYEGQR-DECDXEYFLHEHAPYGGHRTIEKK 297
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 282 VFSRAKRPVIFAPPSW 299
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 298 VFAKAKKRIIFTHPNW 315
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 12
US-10-188-186-20
;; Sequence 20, Application US/10188186
;; Publication No. US20040029789A1
;; GENERAL INFORMATION:
;; APPLICANT: Anderson et al.
;; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;; FILE REFERENCE: 21402-397C
;; CURRENT APPLICATION NUMBER: US/10/188,186
;; PRIOR FILING DATE: 2002-07-02
;; PRIOR APPLICATION NUMBER: 60/303046
;; PRIOR FILING DATE: 2001-07-05
;; PRIOR APPLICATION NUMBER: 60/360814
;; PRIOR FILING DATE: 2002-03-01
;; PRIOR APPLICATION NUMBER: 60/303828
;; PRIOR FILING DATE: 2001-09-07
;; PRIOR APPLICATION NUMBER: 60/323380
;; PRIOR FILING DATE: 2001-09-19
;; PRIOR APPLICATION NUMBER: 60/361133
;; PRIOR FILING DATE: 2002-03-01
;; PRIOR APPLICATION NUMBER: 60/304016
;; PRIOR FILING DATE: 2001-07-09
;; PRIOR APPLICATION NUMBER: 60/304502
;; PRIOR FILING DATE: 2001-07-11
;; PRIOR APPLICATION NUMBER: 60/305262
;; PRIOR FILING DATE: 2001-07-13
;; PRIOR APPLICATION NUMBER: 60/373881
;; PRIOR FILING DATE: 2002-04-19
;; PRIOR APPLICATION NUMBER: 60/305673
;; PRIOR FILING DATE: 2001-07-16
;; Remaining Prior Application data removed - See File Wrapper or PAM.
;; NUMBER OF SEQ ID NOS: 368
;; SOFTWARE: Custom
```

```
;; SEQ ID NO 20
;; LENGTH: 305
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-188-186-20

Query Match      41.7%; Score 677.5; DB 15; Length 305;
Best Local Similarity 49.4%; Pred. No. 1.4e-61; Indels 3; Gaps 3;
Matches 121; Conservative 49; Mismatches 72;

QY 55 HFGSYSSVDDGKPLVREPCRSQAVSSSGQMLGSLGAEIDSACVFRMNAQPTVGFAD 114
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 61 HY-GYINVKTOEPLQD-CDLCAIVNSGQWGVKGNEDRSSCIWRMNNAPTGYEED 118
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 115 VGORSTLRVSHTSVPLLRNYSHPQAKADTLVMWQGRHMDVLGRTYRTLLQTR 174
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 119 VGRMTMIRVSHTSVPLLRNYSHPQAKADTLVMWQGRHMDVLGRTYRTLLQTR 178
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 175 MYPGLOVTFEPRMMAVCDQIFQDETGNRQSGSFLSTGWTMLALECEIIVVGMV 234
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 179 IYPNAQIVTTEKMSYCDGVFKETGDSVLTLSQSSPOLYQSSSYLSTGFTILA 238
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 235 SDSYCREKSHSPVPHYFEKRLDECQMYLAHEQAPRSARHFTIEKAVFSRAKRPVIF 294
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 239 NDTYCKTEGYRKVPYHYEGQR-DECDXEYFLHEHAPYGGHRTIEKVFAPAKAKKRIIF 297
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 295 AHPSW 299
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 298 THPNW 302
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 13
US-10-097-340-189
;; Sequence 189, Application US/10097340
;; Publication No. US20030087250A1
;; GENERAL INFORMATION:
;; APPLICANT: John MONAHAN
;; APPLICANT: Manjula GANNAVAPAPU
;; APPLICANT: Sebastian HOERSCH
;; APPLICANT: Shubhangi KAMATKAR
;; APPLICANT: Steve G. KOVATIS
;; APPLICANT: Rachel E. MEYERS
;; APPLICANT: Michael MORRISSEY
;; APPLICANT: Peter OLANDT
;; APPLICANT: Ami SEN
;; APPLICANT: Peter VEIBY
;; APPLICANT: Gordon B. MILLS
;; APPLICANT: Robert C. BAST, Jr.
;; APPLICANT: Karen LU
;; APPLICANT: Rosemarie SCHMANDT
;; APPLICANT: Xumei ZHAO
;; APPLICANT: Karen GLATT
;; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
;; ASSESSMENT, Prevention, and Therapy of Ovarian Cancer
;; FILE REFERENCE: MRI-030
;; CURRENT APPLICATION NUMBER: US/10/097,340
;; PRIOR FILING DATE: 2002-03-14
;; PRIOR APPLICATION NUMBER: 60/276,025
;; PRIOR FILING DATE: 2001-03-14
;; PRIOR APPLICATION NUMBER: 60/325,149
;; PRIOR FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: 60/276,026
;; PRIOR FILING DATE: 2001-03-14
;; PRIOR APPLICATION NUMBER: 60/324,967
;; PRIOR FILING DATE: 2001/09/26
;; PRIOR APPLICATION NUMBER: 60/311,732
;; PRIOR FILING DATE: 2001-08-10
;; PRIOR APPLICATION NUMBER: 60/325,102
;; PRIOR FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: 60/323,580
;; PRIOR FILING DATE: 2001-09-19
;; NUMBER OF SEQ ID NOS: 363
;; SOFTWARE: FastSeq for Windows Version 4.0
```

SEQ ID NO 189
LENGTH: 336
TYPE: PRT
ORGANISM: Homo sapiens
US-10-097-340-189

Query Match 39.3%; Score 639.5; DB 14; Length 336;
Best Local Similarity 47.9%; Pred. No. 1,4e-57;
Matches 125; Conservative 44; Mismatches 89; Indels 3; Gaps 3;

QY 44 TGSRPVTP-GPLHFGSYSPVDGKPLVREPCSCAVVSSGQMLGSLGAIDSAECVFR 102
DB 64 TQGRPGVAPRPLDGLGVADHKPL-KMHCRCALVTSSGHLHSGSGSIDDTCEVIR 122
QY 103 MNOAPVGFADVGQSTLVVSHTSVPLLRNYSHFQKARDTLVWVGGRHMDRVLG 162
DB 123 MNDAPRTGGRDVGNRSLVIAHSSIQRLIRNHDLLNVSQGTVFLFWGSSYMRDCK 182
QY 163 GRVYRTLLQTRMYPGLOVYTFERRMAYCDQIFQDETGNRRSGSFLSTGWTMLAL 222
DB 183 GQVYNNHLHSQVLPRLKAFMITRHKMLQDFELFKQETGDKRKISNTWLSTGWTMLAL 242
QY 223 ELCEEIVYGVMSDSYCREKSHSPVPHYFEKGRLDCEQMYLAHQAPR-SAHRFITEKA 281
DB 243 ELCDRINVGVMPDPFCRDNNHSPVPHYTFEPGPDCTMYLSHERGRKSHHRFITEKR 302
QY 282 VFSRMAKRPIVFAHPSWRT 302
DB 303 VFKWMARTFNIHFQPDWKPE 323

RESULT 14
US-10-094-749-2441
Sequence 2441, Application US/10094749
Publication No. US20030219741A1

GENERAL INFORMATION:

APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHITO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
PRIOR FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2441
LENGTH: 336
TYPE: PRT
ORGANISM: Homo sapiens
US-10-094-749-2441

Query Match 39.3%; Score 639.5; DB 14; Length 336;
Best Local Similarity 47.9%; Pred. No. 1,4e-57;
Matches 125; Conservative 44; Mismatches 89; Indels 3; Gaps 3;

QY 44 TGSRPVTP-GPLHFGSYSPVDGKPLVREPCSCAVVSSGQMLGSLGAIDSAECVFR 102
DB 64 TQGRPGVAPRPLDGLGVADHKPL-KMHCRCALVTSSGHLHSGSGSIDDTCEVIR 122
QY 103 MNOAPVGFADVGQSTLVVSHTSVPLLRNYSHFQKARDTLVWVGGRHMDRVLG 162
DB 123 MNDAPRTGGRDVGNRSLVIAHSSIQRLIRNHDLLNVSQGTVFLFWGSSYMRDCK 182
QY 163 GRVYRTLLQTRMYPGLOVYTFERRMAYCDQIFQDETGNRRSGSFLSTGWTMLAL 222
DB 183 GQVYNNHLHSQVLPRLKAFMITRHKMLQDFELFKQETGDKRKISNTWLSTGWTMLAL 242
QY 223 ELCEEIVYGVMSDSYCREKSHSPVPHYFEKGRLDCEQMYLAHQAPR-SAHRFITEKA 281
DB 243 ELCDRINVGVMPDPFCRDNNHSPVPHYTFEPGPDCTMYLSHERGRKSHHRFITEKR 302
QY 282 VFSRMAKRPIVFAHPSWRT 302
DB 303 VFKWMARTFNIHFQPDWKPE 323

RESULT 15
US-10-288-252-11
Sequence 11, Application US/10288252
Publication No. US20030143686A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: LAL, Preeti G.
APPLICANT: TANG, Y. Tom
APPLICANT: YUE, Henry
APPLICANT: BURFORD, Neil
APPLICANT: GANDHI, Ameeta R.
APPLICANT: WARREN, Bridget A.
APPLICANT: YAO, Monique G.
APPLICANT: TRIBOUTLEY, Catherine M.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: LEE, Ernestine A.
APPLICANT: HAFALIA, April J.A.
APPLICANT: LU, Yan
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: SANJUNWALA, Madhu S.
APPLICANT: DING, Li
TITLE OF INVENTION: TRANSFERASES
FILE REFERENCE: PI-0241 USA
CURRENT APPLICATION NUMBER: US/10/288,252
PRIOR FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: PCT US01/30424
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US 60/252,819
PRIOR FILING DATE: 2000-11-21
PRIOR APPLICATION NUMBER: US 60/249,639
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: US 60/247,931
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/246,001
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US 60/244,025
PRIOR FILING DATE: 2000-10-27
PRIOR APPLICATION NUMBER: US 60/238,481
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/236,523
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PERL Program
SEQ ID NO 11
LENGTH: 336
TYPE: PRT
ORGANISM: Homo sapiens
US-10-288-252-11

Query Match 39.3%; Score 639.5; DB 14; Length 336;
Best Local Similarity 47.9%; Pred. No. 1,4e-57;
Matches 125; Conservative 44; Mismatches 89; Indels 3; Gaps 3;

Query Match 39.1%; Score 636.5; DB 14; Length 336;
 Best Local Similarity 47.5%; Pred. No. 2,9e-57;
 Matches 124; Conservative 45; Mismatches 89; Indels 3; Gaps 3;

QY	44	TGSRPTVP-GPLHFGSYSSVPDGGKPLVREPCAVVSSGOMLGSLGAELIDSAECVPR	102
DB	64	TQRPQVAPAGPRPLDGLGVADHKPL-KMHCRCALVTSGLHLHSROGQIDQTECVIR	122
QY	103	MNQAPTVGFADYGGRETLVSHTSVPLLRNYSHYFQKARDTLVVMWGQGRHMDRVLG	162
DB	123	MNDAPTKYGRDVGNRITSLRVIAHSSIQRIIRNRHDLINVSOGTVFIWGPSSYMRDGG	182
QY	163	GRTYRTLLQLTRMYPGLQVYTFTERMMAYCDQIFODETGKRRROSGSFLSTGWTMLAL	222
DB	183	GQYYNNLHLISQVLPRLKAFMITRHKQLQFDELFKQETGKORKISNTWLSTGWTMLAL	242
QY	223	ELCEELVYGVMSDSYCREKSHPSVPYHYFEKGRLDCCOMYLAHQAPR-SAHREITEKA	281
DB	243	ELCDRIINVGVVPPDFCRDPNHPSPYHYDPFGPDCTMYLSHERGKSGSHRPLETKR	302
QY	282	VFSRWAKKRPIVFAHPSMRTE	302
DB	303	VFKNMARTFNHIFQPDMPKE	323

Search completed: December 21, 2004, 22:50:58
 Job time : 79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 21, 2004, 22:44:46 ; Search time 105 Seconds
(without alignments)
2044.364 Million cell updates/sec

Title: US-10-019-114A-7
Perfect score: 1626
Sequence: 1 MKRGRVLVILICSVFSAV.....FSRMAKRPIVFAHPSWRT 302

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delcxt 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+p2n.model -DEV=xlh
-Q=/cg2_1/USPTO.spool/US10019114/runat_21122004_100141_3443/app_query.fasta_1.455
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=p2n -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10019114.@CGN_1.1.69@runat_21122004_100141_3443 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NRG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA.*
1: /cg2_6/ptodata/1/ina/5A COMB.seq.*
2: /cg2_6/ptodata/1/ina/5B COMB.seq.*
3: /cg2_6/ptodata/1/ina/6A COMB.seq.*
4: /cg2_6/ptodata/1/ina/6B COMB.seq.*
5: /cg2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cg2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1626	100.0	1556	4	US-09-599-360B-25
2	639.5	39.3	2056	3	US-09-334-601-12
3	544.5	33.5	2697	4	US-09-620-312D-401
4	494	30.4	546	4	US-09-621-976-55
5	196.5	12.1	2026	1	US-08-503-133A-1
6	196.5	12.1	2026	2	US-08-576-775A-1
7	196.5	12.1	2026	2	US-08-972-498-1
8	196.5	12.1	2026	3	US-08-899-545-1
9	194.5	12.0	1020	2	US-08-446-875-15
10	192.5	11.8	1218	2	US-08-446-875-1
11	192.5	11.8	1218	2	US-08-102-385G-1
12	190.5	11.7	1048	1	US-08-626-994A-4

13	190.5	11.7	1048	3	US-08-957-742-4	Sequence 4, Appl
14	190.5	11.7	1660	1	US-08-626-994A-2	Sequence 2, Appl
15	190.5	11.7	1660	3	US-08-957-742-2	Sequence 2, Appl
16	188.5	11.6	1292	1	US-08-503-133A-7	Sequence 7, Appl
17	188.5	11.6	1292	2	US-08-576-775A-7	Sequence 7, Appl
18	188.5	11.6	1292	2	US-08-972-498-7	Sequence 7, Appl
19	188.5	11.6	1292	3	US-08-899-545-7	Sequence 7, Appl
20	188.5	11.6	1682	1	US-08-503-133A-8	Sequence 8, Appl
21	188.5	11.6	1682	2	US-08-576-775A-8	Sequence 8, Appl
22	188.5	11.6	1682	2	US-08-972-498-8	Sequence 8, Appl
23	188.5	11.6	1682	3	US-08-899-545-8	Sequence 8, Appl
24	179	11.0	1188	2	US-08-446-875-9	Sequence 9, Appl
25	179	11.0	1188	2	US-08-102-385G-9	Sequence 9, Appl
26	178	10.9	1206	3	US-08-666-367B-2	Sequence 2, Appl
27	178	10.9	1206	3	US-09-143-438-2	Sequence 2, Appl
28	178	10.9	2671	2	US-08-666-367B-1	Sequence 1, Appl
29	178	10.9	2671	3	US-09-143-438-1	Sequence 1, Appl
30	176	10.8	1206	2	US-08-446-875-3	Sequence 3, Appl
31	176	10.8	1206	2	US-08-102-385G-3	Sequence 3, Appl
32	174.5	10.7	1622	3	US-09-334-601-9	Sequence 9, Appl
33	172	10.6	2117	4	US-08-361-304A-1	Sequence 1, Appl
34	171	10.5	1146	2	US-08-666-367B-4	Sequence 4, Appl
35	171	10.5	1146	3	US-09-143-438-4	Sequence 4, Appl
36	168	10.3	1246	1	US-08-446-777-3	Sequence 3, Appl
37	168	10.3	1540	4	US-09-023-655-1250	Sequence 1250, Ap
38	168	10.3	2304	1	US-08-446-777-5	Sequence 5, Appl
39	168	10.3	2304	1	US-08-446-777-7	Sequence 7, Appl
40	167	10.3	2528	4	US-10-140-002-461	Sequence 461, App
41	150.5	9.3	1158	2	US-08-446-875-11	Sequence 11, Appl
42	150.5	9.3	1158	2	US-08-102-385G-11	Sequence 11, Appl
43	150.5	9.3	1766	1	US-07-991-587A-6	Sequence 6, Appl
44	150.5	9.3	1766	1	US-08-309-985-6	Sequence 6, Appl
45	149	9.2	2121	4	US-09-425-488-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-599-360B-25
; Sequence 25, Application US/09599360B
; Patent No. 6548633
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bouguetier, L.
; TITLE OR INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: GENSET.050CP3
; CURRENT APPLICATION NUMBER: US/09/599,360B
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113,686
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/141,032
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/469,099
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patent.pm
; SEQ ID NO 25
; LENGTH: 1556
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 261..1166
; NAME/KEY: sig_peptide
; LOCATION: 261..314
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 8.80
; NAME/KEY: polyA_site
; LOCATION: 1524..1556
; US-09-599-360B-25

Alignment Scores:

Pred. No.: 3,61e-179 Length: 1556
 Score: 1626.00 Matches: 302
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-019-114a-7 (1-302) x US-09-599-360B-25 (1-1556)

```

QY 1 MetLysAlaProGlyArgLeuValLeuIleLeuCySerValAlaPheSerAlaVal 20
Db ATGAAGGCTCGGGTGGCTCGTGCATCATCTGTGCTCCGGGCTTCTGCGCCGTC 320
QY 21 TyrTlleuLeuCySerTTPAlaGlyLeuProLeuCySerAlaThrCysLeuAspHis 40
Db TACATCTCTGCTGCTGCTGGCGCGCTCCCTGCTGCTGGCCACTCGAGACAC 380
QY 41 HisPheProThrGlySerArgProThrValProGlyProLeuHisPheSerGlyTyrSer 60
Db CACTTCCCAAGGCTTCAGGCCACTGTGCGGAGCCCTTGCACCTTCACTGATATAGC 440
QY 61 SerValProAspGlyLysProLeuValArgGluProCysArgSerCysAlaValSer 80
Db AGTGTCCAGATGGGAAGCCGCTGTGTCGAGAGCCCTGCGCAGCTGTGCGTGTCC 500
QY 81 SerSerGlyGlnMetLeuGlySerGlyLeuGlyValAlaGluIleAspSerAlaGluCysVal 100
Db AGCTCCGCGCAAAATGCTGGGCTCAGGCTGGGTCTAGATCGACAATGCGGAGTGGTG 560
QY 101 PheArgMetAsnGlnAlaProThrValAlaGlyPheGlnAlaAspValAlaGlyIleArgSerThr 120
Db TTCCGCTGTAACACAGGGGCCACCGTGGCTTTGAGGCGGATGTGGGCCACGCGACAGC 620
QY 121 LeuArgValAlaSerHisThrSerValProLeuLeuLeuArgAsnTyrSerHisTyrPhe 140
Db CTCGCTGCTGCTCACACACAGCGTCCGCTGCTGCGCAACTTCACTACTTTC 680
QY 141 GlnLysAlaArgAspThrIleuTyrMetValTyrGlyGlnGlyArgHisSmetAspArgVal 160
Db CAGAAGGCCCGAAGACAGGCTCTACATGATGGGGCCAGGCGAGGACATGAGACCGGGTG 740
QY 161 LeuGlyValArgThrTyrArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGln 180
Db CTGGCGGCGCGCACCTACCGCAGCTGCTGCACTTCCAGAGATGACCCCGGCTGAG 800
QY 181 ValTyrThrPheThrGluArgMetLeuAlaTyrCysAspGlnIlePheGlnAspGluThr 200
Db GTGTAACCTTTCAGGAGCGCATGATGGCTTACGACCAAGATCTTCCAGAGACGAGACG 860
QY 201 GlyLysAsnArgArgGlnSerGlySerPheLeuSerThrGlyTyrPheThrMetIleLeu 220
Db GGCAAGAACCGGAGCGAGTGGGCTCTCTTCAGCACCGGCTGTTCACCATGATCTTC 920
QY 221 AlaLeuGlnLeuCySerGluGluIleValIleValTyrGlyMetValSerAspSerTyrCysArg 240
Db GCCTGAGAGCTGTGTAGAGATGTGTCTATAGGATGTGTCAAGCAGCACTACTGAGG 980
QY 241 GlnLysSerHisProSerValProTyrHisTyrPheGlnLysGlyArgLeuAspGluCys 260
Db GAGAAGAGCACCCCTCAGTGCTTACCACTTATAGAAAGGCGGCTAGATGAGTGT 1040
QY 261 GlnMetTyrLeuAlaHisGluGlnAlaProArgSerAlaHisArgPheIleThrGluLys 280
Db GAGATGATCTGAGACAGCAGAGCGCCCGGAGCGCCCACTTCACTACAGAGAG 1100
QY 281 AlaValPheSerArgTTPAlaLysLysArgProIleValPheAlaHisProSerTyrArg 300
Db GCGGCTCTTCTCCCTGCGTGGCCAGAAAGGCCCATCTGTTCGCCCATCTCGTGGAGG 1160
QY 301 ThrGlu 302
Db 1161 ACTGAG 1166

```

RESULT 2

US-09-334-601-12

; Sequence 12, Application US/09334601
 ; Patent No. 6280989
 ; GENERAL INFORMATION:
 ; APPLICANT: Kapilnov, Dmitri
 ; APPLICANT: Yu. Robert
 ; TITLE OF INVENTION: NOVEL SIALYLTRANSFERASES
 ; FILE REFERENCE: VCUIP-6
 ; CURRENT APPLICATION NUMBER: US/09/334,601
 ; CURRENT FILING DATE: 1999-06-17
 ; NUMBER OF SEQ ID NOS: 94
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 12
 ; LENGTH: 2056
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (181)..(1188)
 US-09-334-601-12

Alignment Scores:

Pred. No.: 4.01e-64 Length: 2056
 Score: 639.50 Matches: 125
 Percent Similarity: 64.75% Conservative: 44
 Best Local Similarity: 47.89% Mismatches: 89
 Query Match: 39.33% Indels: 3
 DB: 3 Gaps: 3

US-10-019-114a-7 (1-302) x US-09-334-601-12 (1-2056)

```

QY 44 ThrGlySerArgProThrValPro---GlyProLeuHisPheSerGlyTyrSerSerVal 62
Db ACCGAGAGGCGCCCGGGTCCCGCGGAGCGGCCACTGACGAGATCTCGGAGTG 429
QY 63 ProAspGlyLysProLeuValArgGluProCysArgSerCysAlaValAlaSerSerSer 82
Db GCGGACCAAGCCCTG---AAATGCACTGCAAGGAGCTGTGCTCGTGACAGCTCA 486
QY 83 GlyGlnMetLeuGlySerGlyLeuGlyAlaGluIleAspSerAlaGluCysValPheArg 102
Db GGCATCTGTGCAAGTGGCAAGTCCCAATGACCAAGAGTGTGTCATCCGC 546
QY 103 MetAsnGlnAlaProThrValGlyPheGlnAlaAspValAlaGlnArgSerThrLeuArg 122
Db ATGAATGACCGCCCAACCGGCTATGGCGTGAAGCGGCAATGCGACACGCTGAGG 606
QY 123 ValValSerHisThrSerValProLeuLeuLeuArgAsnTyrSerHisTyrPheGlnLys 142
Db GTCATGCGCATTTCCAGATCCAGATCCAGAGATCTTCCGACCGCATGACTGTCACAGTG 666
QY 143 AlaArgAspThrLeuTyrMetValTyrGlyGlnGlyArgHisMetAspArgValLeuGly 162
Db AGCGAGGAGCACCGTTCATCTTCTGGGCGCCAGCAGTACATGCGGCGGAGCGGCAAG 726
QY 163 GlyArgThrTyrArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyr 182
Db GGCCAGGCTTACAAACAACCTGCATCTCTGAGCAGAGTGTGCCCCGCGTGAAGGCTTC 786
QY 183 ThrPheThrGluArgMetMetAlaTyrCysAspGlnIlePheGlnAspGluThrGlyLys 202
Db ATGATTAATCTGCCAACAAGATGCTGCACTTTGATGAGCTTTTCAAGCAGGAGACTGGCAA 846
QY 203 AsnArgArgGlnSerGlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeu 222
Db GACAGAGAAATATCCAACACTTGGCTCAGACCTGGCGTTCATTAATGACAAATGCACTG 906
QY 223 GlnLeuCySerGluGluIleValValTyrGlyMetValSerAspSerTyrCysArgGluLys 242
Db GAGCTGTGACAGATCAATGTTATGCGATGTGCCCCAGACTTTCGAGGATCCC 966

```

```

QY 243 SerHisProSerValProTyrHisTyrPheGluIuysglYargLeuAspGluCysGlnMet 262
Db 967 AATACCTTCACTTACCTTATCATATATATGAACTTTGGACCTGATGAAATGTAACAATG 1026
QY 263 TyrLeuAlaHisGluGlnAlaProArg---SerAlaHisArgPheIleThrGluIuysAla 281
Db 1027 TACTCTCCCATGAGCAGAGCAGCAGGAGGAGTATCACCCTTTATCAGAGAAACGA 1086
QY 282 ValPheSerArgTyrAlaIuysLysArgProIleValPheAlaHisProSerTyrArgThr 301
Db 1087 GTCTTTAAGACATGGGACGAGCATTCATATTCATCTTTTCAACCACTGGAAACA 1146
QY 302 Glu 302
Db 1147 GAA 1149

RESULT 3
US-09-620-312D-401
; Sequence 401, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghaast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 401
; LENGTH: 2697
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)..(1136)
US-09-620-312D-401

Alignment Scores:
Pred. No.: 7,13e-53 Length: 2697
Score: 544.50 Matches: 115
Percent Similarity: 57.86% Conservative: 47
Best Local Similarity: 41.07% Mismatches: 102
Query Match: 33.49% Indels: 16
DB: 4 Gaps: 6

US-10-019-114A-7 (1-302) x US-09-620-312D-401 (1-2697)
QY 26 CysTyrAlaGlyLeuProLeuCysLeuAlaThrCysLeuAspHisHisPheProThrGly 45
Db 323 TGCCATAGAGGTCTTCATACGCGTC-----CTGCGCGG 358
QY 46 SerArgProThrValProGlyProLeu-----HisPhe-SerGlyTyrSerSerVa 62

```

```

Db 359 CCGTAGCCCGGACCTGTCAACCTCAAGAAATGAGATCACTACGCTATGTCCCAT 418
QY 62 LProAspGlyLysProLeuValArgIuProCysArgSerCysAlaValAlaSerSerSe 82
Db 419 TCTGGCAACAAGACTGCGCCTCTGG---TGCACACAGTGTGTATGTTCACAGTCTC 475
QY 82 rGlyGlnMetLeuGlySerGlyLeuGlyAlaGluIleAspSerAlaGluCysValPheAr 102
Db 476 CAGCACTCTCTGGGACCAAGCTGCGCCTGTAGATGACCGGCTCAGGTATCAATCTCG 535
QY 102 gMetAsnGlnAlaProThrValGlyPheGluAlaAspValGlyGlnArgSerThrLeuAr 122
Db 536 CATGAATGATGACCAACCACTGCTACTGATGTATGAGCAACAAGACCACTTCCG 595
QY 122 gValAlaSerHisThrSerValProLeuLeuLeuArgAsnTyrSerHisTyrPheGlnIu 142
Db 596 CGTGTGCGCCCATTCAGTGTGTTCGCGGTGTGAGAGAGGCCCAAGAGTTGTCAACCG 655
QY 142 sAlaArgAspThrLeuTyrMetValTyrGlyGlnGlyArgHisMetAspArgValLeuG 162
Db 656 GACCCCTGAAACCGTGTTCATCTTCTGGGGGCCCCCGACGAAAGATGCAAGACCCAGGG 715
QY 162 yGlyArgThrTyrArgThrLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyr 182
Db 716 CAGC---CTGCTGTGTGATCCAGCGAGCGGCTGTGTTCCTCCCAATGGAAGCATA 772
QY 182 rThrPheThrGluArgMetMetAlaTyrCysAspGlnIlePheGlnAspGluThrGlyLy 202
Db 773 TGCCGTCTCCCGGCGGAGTGGGCAATTGAGCACTCTTCGCGGGTGAAGCGGCA 832
QY 202 sAsnArgArgGlnSerGlySerPheLeuSerThrGlyTyrPheThrMetLeuAlaLe 222
Db 833 GGAAGGAGAAAGTCTCATTCGTGTGACACAGGCTGTGTTACCACTGGTATCGCGGT 892
QY 222 uGluLeuCysGluGlnIleValValTyrGlyMetValSerAspSerTyrCysArgGly 242
Db 893 GAGATGTGTGACACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 952
QY 242 sSerHis---ProSerValProTyrHisTyrPheGluIuysglYargLeuAspGluCysG 261
Db 953 GCCCGGCTCCAGCATGCGCTTACCACTACTAGAGCCCAAGGCGGAGCAATGTGT 1012
QY 261 nMetTyrLeuAlaHisGluGlnAlaProArg---SerAlaHisArgPheIleThrGly 280
Db 1013 CACCTACATCAGAAATGAGCAGATGCGCAAGGCAACCAACCGCTTCAACCGAGAA 1072
QY 280 sAlaValPheSerArgTyrAlaIuysLysArgProIleValPheAlaHisProSerTyr 299
Db 1073 AAGGCTCTCTCATCGTGGGCCAGCTGTATGAGCATCACTTCCACCCCTCTCTG 1130

RESULT 4
US-09-621-976-55
; Sequence 55, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 55
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 260..544
; NAME/KEY: sig_peptide

```


RESULT 6
US-08-576-775A-1
Sequence 1, Application US/08576775A
Patent No. 5849904
GENERAL INFORMATION:
APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;
TITLE OF INVENTION: Isolated Polysialyl Transferases,
TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/576, 775A
FILING DATE: 21-December-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/503,133
FILING DATE: 17-July-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/04289
FILING DATE: 22-December-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5849904man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: BOER 1050.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2026 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 301..1377
US-08-576-775A-1

Alignment Scores:
Pred. No.: 1.47e-12 Length: 2026
Score: 196.50 Matches: 71
Percent Similarity: 45.93% Conservative: 42
Best Local Similarity: 28.86% Mismatches: 99
Query Match: 12.08% Indels: 35
Gaps: 8
US-10-019-114A-7 (1-302) x US-08-576-775A-1 (1-2026)
QY 59 TysrSerserValProaspGlyLysProLeuValArgLubProCysArgSerCysAlaVal 78
Db 673 CACAGCCTCTGCTGTAAGTTTCAACCAATGAAAAACCGCAGGTTTAAGACTGCTGCTT 732
QY 79 ValSerSerserGlyLysMetLeuGlySerGlyLeuGlyAlaGluLysSerAlaGlu 98
Db 733 GTTGAAACTCTGCACTTCTACTAGACAGTGATGTGCAAGGATGACAGTCAAT 792
QY 99 CysValPheArgMetLeuGlnAlaProThrValGlyPheGluAlaAspValGlyGlnArg 118

Db 793 TTTGTAATCAGTGCATCTTACGTCCTGTGTGAGAGTTTGTGCGGATGTGGCACTTAA 852
QY 119 SerThrLeuArgValAlaSerHisThrSerValProLeuLeu-----LeuArgAsn 135
Db 853 TCAGATTTTATTACCATGAAACCCATGAGTGTGTGAGAGAGCATTTGGAGCTTCCGAT 912
QY 136 TysrSer-----HisThrPheGlnLysAlaArgPheThrLeu----- 147
Db 913 GAGAGTGACAGACGAAATTTGTGCTATGACTTTCCATGCTGAATGACAGTGTCTTTGG 972
QY 148 -----TyrMetValITrpGlyGlnGlyArgHisMetAspArgValLeuGlyArg 164
Db 973 ATCCCGCTTTCATGCTCAAGAGGAGAGAGACGCTGGAATGGTT---AATGCATTA 1029
QY 165 ThrTyArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyrPhe 184
Db 1030 ATCCTTAAGAAACAGCTGAAGAGTGAAGTGCATCATCATCATGAGACT-TATTCAATC 1088
QY 185 ThrGluArgMetMetAlaTyrCysAspGlnIlePheGlnAspGluThrGlyLysAsn-Arg 204
Db 1089 TGTCAGAGGTTTACGTGCT-----GACCAACAAGTGCCTCATCA 1127
QY 204 GATGlnSerGlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeuGlu 224
Db 1128 AAGACCC-----AGCACAGGCTCTCTCATATGACACACTGACCCAGATT 1172
QY 224 uCysGluGluIleValValTyrGlyMetValSerAspSerTyrCysArgGlyLysSerHis 244
Db 1173 TTGTGATGAATTTACCTGATGATGTTCTGG-----CCCTTCCTTAAGATTTGAATGC 1226
QY 244 sProSerValProTyrHisTyrPheGluLysGlyArgLeuAspGluCysGlnMetTyrLe 264
Db 1227 AAAAGCTGTGAATATCATTTACTACGAT-----GACTTGAATATGATACTT 1274
QY 264 uAlaHisGluGlnAlaProArgSerAlaHisArgPheIleThrGluLysAlaValPhe 284
Db 1275 TTCCACGCCAAGCCCTCAGCAATGCAATGAATGAATGAATGAATGCTGACACAA 1334
QY 284 rArgTrpAlaLysLys 289
Db 1335 CACAGCAGCACTTAA 1350

RESULT 7
US-08-972-498-1
Sequence 1, Application US/08972498
Patent No. 5959078
GENERAL INFORMATION:
APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;
TITLE OF INVENTION: Isolated Polysialyl Transferases,
TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/972, 498
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/576, 775
FILING DATE: 21-December-1995
APPLICATION NUMBER: 08/503,133

Db 673 CACAGCCTCTGCGTGAAGTTTCAACCAATGAAAAACCGCAGTTTAAGACTGCTGTT 732
 Qy 79 ValSerSerGlyGlnMetLeuGlySerGlyLeuGlyValAgluIleAspSerAlaGlu 98
 Db 733 GTTGGAACCTGGCACTTCTACTAGACAGTGAATGTGGCAAGAGATTGACAGTCAACAT 792
 Qy 99 CysValPheArgMetLeuGlnAlaProThrValGlyPheGluAlaAspValGlyGlnArg 118
 Db 793 TTGTATATCAGGGTGCATCTAGCTCCCTGTGTGAGTTTCTCGAGATGTGGGAGACTAAA 852
 Qy 119 SerThrLeuArgValValSerHisThrSerValProLeuLeu-----LeuArgAsn 135
 Db 853 TCGATTATTTATACATGAACCCATCGTGTGACAGAGACTTTGAGGCTTTCCGAAAT 912
 Qy 136 TyrSer-----TyrMetValTTPGlyGlnGlyArgHisMetLeuAspArgValIleuGlyGlyArg 147
 Db 913 GAGAGTACAGACAAATTTGTGCATATAGACTTTCATGCTGATGATGACAGTGTCTTTGG 972
 Qy 148 -----TyrMetValTTPGlyGlnGlyArgHisMetLeuAspArgValIleuGlyGlyArg 164
 Db 973 ATCCCGCTTTCATGTGCAAG 1029
 Qy 165 ThrTyrArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyrThrPhe 184
 Db 1030 ATCTTAAAGAACAGCTGAAGAGTGCAGACCTGCTATTCATCAGTGAAGT-TATTCAATGC 1088
 Qy 185 ThrGluArgMetLeuAlaTyrCysAspGlnIlePheGlnAspGluThrGlyValAsn-Ar 204
 Db 1089 TGTCAAGAGTGTACTGCT-----GACCAACAAAGTGGCCATCA 1127
 Qy 204 GAGGlnSerGlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeuGluLe 224
 Db 1128 AAGAGCC-----AGCAGAGGCTCTCTCATGTATACACACTGCCACACAGATT 1172
 Qy 224 uCysGluGluIleValValTyrGlyMetValSerAspSerTyrCysArgGluIlySerHis 244
 Db 1173 TTGTGATGAATTCACCTGTAATGGGTTCTGG-----CCCTTCCCTTAAGATTGAAATGG 1226
 Qy 244 sProSerValProTyrHisTyrPheGlnIlyValArgLeuAspGluCysGlnMetTyrIle 264
 Db 1227 AAAAGCTGAATATCATCTATCAAGAT-----GACTGTAATATATGATATCTT 1274
 Qy 264 uAlaHisGluGlnAlaProArgSerAlaHisArgPheIleThrGluValAlaValPheSe 284
 Db 1275 TTCACAGCAGAGCCCTCACAGATGTCATTAGATCAAAACCTGTAATGTGTACACAA 1334
 Qy 284 rArgTTPAlaIlyValys 289
 Db 1335 CAGAGCAGCAGCTAAAA 1350
 RESULT 9
 US-08-446-875-15
 ; Sequence 15, Application US/08446875
 ; Patent No. 5858751
 ; GENERAL INFORMATION:
 ; APPLICANT: Paulson, James C.
 ; APPLICANT: Wen, Xiaohong
 ; APPLICANT: Livingston, Brian Duane
 ; APPLICANT: Gillespie, William
 ; APPLICANT: Keim, Sarge
 ; APPLICANT: Burlingame, Alma L.
 ; APPLICANT: Medzhradszky, Katalin
 ; TITLE OF INVENTION: Compositions and Methods for the
 ; TITLE OF INVENTION: Identification and Synthesis of Sialyltransferases
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSES: Pome, Smith, Lande & Rose
 ; STREET: 2029 Century Park East, Suite 3800
 ; CITY: Los Angeles
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 90067
 ; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/446,875
 FILING DATE: July 12, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/102,385
 FILING DATE: August 4, 1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Oldenkamp, David J.
 REGISTRATION NUMBER: 29,421
 REFERENCE/DOCKET NUMBER: 111-197
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (310) 788-5000
 TELEFAX: (310) 277-1297
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1020 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHEICAL: NO
 US-08-446-875-15
 Alignment Scores:
 Pred. No.: 8,36e-13 Length: 1020
 Score: 194.50 Matches: 56
 Percent Similarity: 44.79% Conservative: 30
 Best Local Similarity: 29.17% Mismatches: 67
 Query Match: 11.96% Indels: 39
 DB: 2 Gaps: 6
 US-10-019-114a-7 (1-302) x US-08-446-875-15 (1-1020)
 Qy 73 CysArgSerCysAlaValValSerSerSerGlyGlnMetLeuGlySerGlyLeuGlyAla 92
 Db 415 TGCAGGCGGTGCGCCCGTGTGTGGGCAACTGCGGCACTGAGGAGTCTTTATGGGCTT 474
 Qy 93 GluIleAspSerAlaGluCysValPheArgMetLeuAlaProThrValGlyPheGlu 112
 Db 475 GAGATGACAGTACACACTTGTCTCAGATGAAACAGCGCCACCGCAGGAGTTGAA 534
 Qy 113 AlaAspValGlyGlnArgSerThrLeuArgValValSerHisThrSerValProLeuLeu 132
 Db 535 GCTGATGTTGGAGCCAGACCAACCACTGATGATCCCTGAGAGCTTCCGGAGCTG 594
 Qy 133 LeuArgAsnTyrSerHisTyrPheGlnIlyValArgAspThrLeuTyrMetValTTPGly 152
 Db 595 CCACATATATGTACATGATTCCTGTCGCTTCAAG--ACATGCACTTGAGTGG-- 648
 Qy 153 GlnGlyArgHisMetLeuAspArgValIleuGlyGlyArgThrTyrArgThrLeuLeu----- 170
 Db 649 -----GTGGTGAAGCGCCATCACACGCGGACCATTTCCACACTTACATCCCGGTT 699
 Qy 171 -----GlnLeuThrArgMetTyrProGlyLeuGlnVal 181
 Db 700 CCGCAAGATCAGAGTGAACAGATATGATCTGATCTACCAACCCAGCTTCATCAAG 759
 Qy 182 TyrThrPheThrGluArgMetLeuAlaTyrCysAspGlnIlePheGlnAspGluThrGly 201
 Db 760 TATGCTTTGACAAACGCGTGCAGAGGAC----- 769
 Qy 202 LysAsnArgArgGlnSerGlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAla 221
 Db 790 -----GGGCGATACCATCTACCGGATCTCTCCGTCATCTCTCA 831
 Qy 222 LeuGluLeuCysGluGluIleValValTyrGlyMetValSerAspSerTyrCysArgGlu 241
 Db 832 ATCATGTCTGCGATAGGATGAGCTTACGAGCTTGGGGGACAGAC-----AAAGG 885

QY 242 LysSerHisProSerValProTyrHisTyrPheGlu 253
 Db 886 AACGCGAC-----CACTACTGGAG 906

RESULT 10
 US-08-446-875-1
 Sequence 1, Application US/08446875
 Patent No. 5858751
 GENERAL INFORMATION:
 APPLICANT: Paulson, James C.
 APPLICANT: Men, Xiaohong
 APPLICANT: Livingston, Brian Duane
 APPLICANT: Gillespie, William
 APPLICANT: Kelm, Sarge
 APPLICANT: Burlingame, Alma L.
 APPLICANT: Medzhiradzky, Katalin
 TITLE OF INVENTION: Compositions and Methods for the
 TITLE OF INVENTION: Identification and Synthesis of Sialyltransferases
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Poms, Smith, Lande & Rose
 STREET: 2029 Century Park East, Suite 3800
 CITY: Los Angeles
 STATE: CA
 COUNTRY: USA
 ZIP: 90067
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/446,875
 FILING DATE: July 12, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/102,385
 FILING DATE: August 4, 1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Oldenkamp, David J.
 REGISTRATION NUMBER: 29,421
 REFERENCE/DOCKET NUMBER: 111-197
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (310) 788-5000
 TELEFAX: (310) 277-1297
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1218 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: porcine
 TISSUE TYPE: liver, submaxillary glands
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 91..1119
 OTHER INFORMATION: /product="porcine Gal Beta 1,3
 OTHER INFORMATION: GAlNAc alpha 2,3 sialyltransferase"
 US-08-446-875-1

Alignment Scores:
 Pred. No.: 1.9e-12 Length: 1218
 Score: 192.50 Matches: 64
 Percent Similarity: 46.79% Conservative: 38
 Best Local Similarity: 29.36% Mismatches: 82
 Query Match: 11.84% Indels: 34
 DB: 2 Gaps: 7

US-10-019-114a-7 (1-302) x US-08-446-875-1 (1-1218)

QY 40 HisHisPheProThrGlySerArgPro-ThrValProGlyProLeuHisPheSerGlyTyr 59
 Db 444 CACCAT-----CAGGAGCTGTTCCAGGTGGTGGTGGAGCTG----- 483

QY 59 rSerSerValProAspGlyLysProLeuValArgGlu-----ProCysArgSerGly 76
 Db 484 -----GACCCCTGCTGGAGAAAGAGCTGTGTCAGCTGCCGGCGCTG 524

QY 76 bAlaValIalSerSerSerGlyGlnMetLeuGlySerGlyLeuGlyAlaGluIleAsp 96
 Db 525 CGCGCTGTCGTGCTGAGTGAACAAGGCTCCACGAGAGGGCTTGAAGGCCCTCAAGACAG 584

QY 96 rAlaGluCysValPheArgMetAsnGlnAlaProThrValGlyPheGluIleAspValG 116
 Db 585 CCAAGACTTCGTGCTGAGTGAACAAGGCTCCACGAGAGGGCTTGAAGGCCCTCAAGACAG 644

QY 116 yGlnArgSerThrLeuArgValAlaSerHisThrSerValProLeuLeuValArgAsnTyr 136
 Db 645 GAGCAAGACCAACCACCATTTGTCGTAACCCGAGAGCTTCCGGAGCTGGCGAGAGAGCT 704

QY 136 rSerHisTyrPheGlnLysAlaArgAspThrLeuTyrMetValTTPGlyGlnGlyArgH 156
 Db 705 CAGC-----ATGATCTGCTGCTCCCTTCAAGACCAACCGA 737

QY 156 sMetAspArgValLeuGlyGlyArgThrTyrArgThrLeuLeuGlnLeuThrArgMetTyr 176
 Db 738 CCTGAGGTGGTATGATGACCCGACCAACCCGACCATTCACACCACTTACCTCTGT 797

QY 176 rProGlyLeuGlnValTyrThrPheThrGluArgMetMetAlaTyrCysAspGlnIlePh 196
 Db 798 CCCCCC---AAATCAAACTCAAAAAGAGAAAGATTCCTGATTAACCCGGCCTTCAT 854

QY 196 eGlnAspGluThrGlyLysAsnArgArgGlnSerGlySerPheLeuSerThrGlyTTPH 216
 Db 855 CAAGTACGTCTTCGACAGGTGCTGACAGGCCACGGCGCTGACCCGCTCACTGCAATCCT 914

QY 216 eThrMetIleLeuAlaLeuGluLeuGluGlnGluIleValTyrGlyMetValSerAs 236
 Db 915 CTCGTGATCTTCTCCCTGCAATCTGTGACGAGGTGAGCACTTGATAGCTTTGGGGCGGA 974

QY 236 pSerTyrCysArgGluLysSerHisProSerValProTyrHisTyrPheGlu 253
 Db 975 CAGC-----AAGGGAACGCGAC-----CACTACTGGAG 1005

RESULT 11
 US-08-102-385G-1
 Sequence 1, Application US/08102385G
 Patent No. 5962294
 GENERAL INFORMATION:
 APPLICANT: Paulson, James C.
 APPLICANT: Men, Xiaohong
 APPLICANT: Livingston, Brian Duane
 APPLICANT: Gillespie, William
 APPLICANT: Kelm, Sarge
 APPLICANT: Burlingame, Alma L.
 APPLICANT: Medzhiradzky, Katalin
 TITLE OF INVENTION: Compositions and Methods for the
 TITLE OF INVENTION: Identification and Synthesis of Sialyltransferases
 NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
 STREET: 2029 Century Park East, 38th Floor
 CITY: Los Angeles
 STATE: CA
 COUNTRY: USA
 ZIP: 90067
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/102,385G
FILING DATE: 04-AUG-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/925369
FILING DATE: 04-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Oldenkamp, David J.
REGISTRATION NUMBER: 29,421
REFERENCE/DOCKET NUMBER: 97-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310)788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEICAL: NO
ORGANISM: porcine
TISSUE TYPE: liver, submaxillary glands
FEATURE:
NAME/KEY: CDS
LOCATION: 91..119
OTHER INFORMATION: /product= "porcine Gal Beta 1,3
OTHER INFORMATION: GalNAc alpha 2,3 sialyltransferase"
US-08-102-385G-1

Alignment Scores:
Pred. No.: 1.9e-12 Length: 1218
Score: 192.50 Matches: 64
Percent Similarity: 46.79% Conservative: 38
Best Local Similarity: 29.36% Mismatches: 82
Query Match: 11.84% Indels: 34
DB: 2 Gaps: 7

US-10-019-114A-7 (1-302) x US-08-102-385G-1 (1-1218)
QY 40 HsHhPheProHrglySerArgPro-ThrValProGlyProLeuHhPheSerGlyTy 59
Db 444 CACCAT-----CAGGAGCTGTTCAGGTGCTCGGAGACGTG----- 483
QY 59 rSerSerValProApGlyLysProLeuValArgGlu-----ProCysArgSerCy 76
Db 484 -----GACCCCTGCTGAGAAAGAGCTGCTGCGCGCTG 524
QY 76 sAlaValValSerSerSerGlyGlnMetLeuGlySerGlyLeuGlyAlaGluIleAps 96
Db 525 CCGCGTCTGTGGGCACTCGGCAACCTGAAGAGTCTTACTATGCGGCTCAGATAGACAG 584
QY 96 rAlaGluCysValPheArgMetAanglnAlaProThrValGlyPheGlnAlaApValG 116
Db 585 CACGACTTCTGCTGAGTGAACCAAGGCCCCCAGAGGAGGTTTGAAGCCGACGTCGG 644
QY 116 yGlnArgSerThrLeuArgValValSerHhThrSerValProLeuLeuArgAenTy 136
Db 645 GACCAAGACACACCATTCCTGTCATCCCGAGAGCTTCCGGAGCTGCGCAGAGAGT 704
QY 136 rSerHhTyPheGlnLysAlaArgAphThrLeuTyMetValITrpGlyGlnGlyArgH 156
Db 705 CAGC-----ATGATCTGTCCTCCCTTCAAGACACCGA 737
QY 156 sMetAspArgValLeuGlyGlyArgThrTyArgThrLeuGlnLeuHhArgMetTy 176
Db 738 CCGGAGTGTGATTCAGCCACACACCGGACCATCTCCACACCTAAGTTCTGT 797
QY 176 rProGlyLeuGlnValTyThrPheThrGluArgMetMetAlaTyCysAspGlnIlePh 196
Db 798 CCCCCG---AAGATCAAAAGTCAAAAAGAGAAAGATCTGATTATACCCGCGCTTCAT 854

QY 196 eGlnAspGluThrGlyLysAangArgGlnSerGlySerPheLeuSerHrglyTrpH 216
Db 855 CAAATAGTCTTCGACGAGGCTGACGAGGCGACCGGCTGCTGCTGCTGCTGCT 914
QY 216 eTrnMetIleLeuAlaLeuGluLeuCysGluGluIleValValTyArgIleValSerAs 236
Db 915 CTCGATCTTCTTCCTGACATCTGTGACGAGTGTGACTGTGATGCTTTGGGCGGA 974
QY 236 pSerTyCysArgGluLysSerHhProSerValProTyHhTyPheGlu 253
Db 975 CAGC-----AAAGGAATGCGAC-----CACTAGTGGAG 1005

RESULT 12
US-08-626-994A-4
Sequence 4, Application US/08626994A
Patent No. 5798244
GENERAL INFORMATION:
APPLICANT: Shuichi TSUJI et al.
TITLE OF INVENTION: Sia'2,3Gal'1,4GlcNAc '2,8-
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenblum & Bernstein, P.L.C.
STREET: 1941 Roland Clarke Place
CITY: Reston
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 20191
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 diskette
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Perfect 5.1+ (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/626,994A
FILING DATE: April 3, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 77469/1995
FILING DATE: April 3, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Arnold Turk
REGISTRATION NUMBER: 33,094
REFERENCE/DOCKET NUMBER: P14595
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 716-1191
TELEFAX: (703) 716-1180
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1048 bases
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
US-08-626-994A-4
Alignment Scores:
Pred. No.: 2.55e-12 Length: 1048
Score: 190.50 Matches: 65
Percent Similarity: 44.29% Conservative: 28
Best Local Similarity: 30.95% Mismatches: 88
Query Match: 11.72% Indels: 29
DB: 1 Gaps: 7
US-10-019-114A-7 (1-302) x US-08-626-994A-4 (1-1048)
QY 60 SerSerValProApGlyLysProLeuValArgGluProCysArgSerCysAlaValVal 79
Db 326 TCCCTGCTCCCAAGATGTGCGCCCATTAATGAAGCTTAATGTTTGTGCTGTG 385
QY 80 SerSerSerGlyGlnMetLeuGlySerGlyLeuGlyAlaGluIleAspSerAlaGluCys 99
Db 386 GAAAGAGTGTGATCTTGAACAGGAGTCACTGTGCAAGAAATATGAATCAATCATTTT 445

STREET: 1941 Roland Clarke Place
CITY: Reeston
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 20191
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 diskette
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Perfect 5.1+ (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/626,994A
FILING DATE: April 3, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 77469/1995
FILING DATE: April 3, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Arnold Turk
REGISTRATION NUMBER: 33,094
REFERENCE/DOCKET NUMBER: P14595
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 716-1191
TELEFAX: (703) 716-1180
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1660 bases
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
US-08-626-994A-2

Alignment Scores:
Pred. No.: 5,34e-12 Length: 1660
Score: 190.50 Matches: 65
Percent Similarity: 44.29% Conservative: 28
Best Local Similarity: 30.95% Mismatches: 88
Query Match: 11.72% Indels: 29
DB: Gaps: 7

US-10-019-114A-7 (1-302) x US-08-626-994A-2 (1-1660)

QY 60 SerSerValProaspGlyLysProLeuValArgLysProCysArgSerCysAlaValVal 79
DB 510 TCCCTGCTCCCAATGTCGCCCATTAATGAATTAATGTTTGGCTGTGTT 569
QY 80 SerSerSerGlyGlnMetLeuGlySerGlyLeuGlyAlaGluLeuLeuSerAlaGluCys 99
DB 570 GGAACAGTGGATCTTGACAGGAGTCACTGTGACAAAGAAATGATTAATCGATTTT 629
QY 100 ValPheArgMetArgGlnAlaProThrValGlyPheGluAlaAspValGlyGlnArgSer 119
DB 630 GTTTCGATGACATTTTGCCTCCGACAGAGCTTTCCACAAAGATGTTGAGAGGAAAAAC 689
QY 120 ThrLeuArgValValSerHisThrSerValProLeuLeuLeuArgSerHisThr 139
DB 690 AACCTGACA-----ACCTTCATCCGAGCATTTAGAGAAATATTAACAACAT 737
QY 140 PheGlnLysAlaArgAspThrLeuThrMetValTrpGlyGlnGlyArgHisMetAspArg 159
DB 738 CTTTAAACCATTCAGACACCTTAACAATTCTTCTCCTCACTTTAAAAAGCTTGATGGGCC 797
QY 160 ValLeu-----GlyGlyArgThrTrpArgThrLeuLeu 170
DB 798 ATACTTGTGATCCCTGATTTTCTTCACACATCTTCGCACTGTAAACGAGAGCTAGTG 857
QY 171 GlnLeuThrArgMetTrpArgLys--LeuGlnValArgThrPheThrGlnArgMetMet 189
DB 858 GATTTTGTGTTGAGACAGAGGTCAGTAAAGCTC-----CACTTG 899
QY 190 AlaTrpCysAspGlnLeuPheGln-----AspGluThrGlyLysAsnArgSerGlnSer 207
DB 900 GCTTGGCCTGGAATATCATGCAACATGTCAACACAGTATCGAAAAACAAACACCTGTCA 959

QY 208 GlySerPheLeuSerThrGlyTrpPheThrMetLeuAlaLeuGluLeuCysGluGlu 227
DB 960 CCCAAGACTGAGACAGATATCTTAATGTATATCTTGATCTGCAATATGTGAAGAG 1019
QY 228 IleValValTrpGlyMetValSer-----AspSerTrpCysArgGluLysSerHis 244
DB 1020 ATCCACTGTGACGGTTTCTGCGCCTTTGGATTGTGACCCCAACACAGGAG----- 1070
QY 245 ProSerValProTrpHisTrpPheGluLys 254
DB 1071 ---GATCTGCCCTACCACTACTATGACAAA 1097

RESULT 15
US-08-957-742-2
Sequence 2, Application US/08957742
Patent No. 6017743
GENERAL INFORMATION:
APPLICANT: Shuichi TSUJI et al.
TITLE OF INVENTION: Sia², 2,3Ga1a 1,4GlcNAc², 2,8-
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenblum & Bernstein, P.L.C.
STREET: 1941 Roland Clarke Place
CITY: Reeston
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 20191
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 diskette
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Perfect 5.1+ (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/957,742
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/626,994
FILING DATE: April 3, 1996
APPLICATION NUMBER: JP 77469/1995
FILING DATE: April 3, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Arnold Turk
REGISTRATION NUMBER: 33,094
REFERENCE/DOCKET NUMBER: P14595
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 716-1191
TELEFAX: (703) 716-1180
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1660 bases
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
US-08-957-742-2

Alignment Scores:
Pred. No.: 5,34e-12 Length: 1660
Score: 190.50 Matches: 65
Percent Similarity: 44.29% Conservative: 28
Best Local Similarity: 30.95% Mismatches: 88
Query Match: 11.72% Indels: 29
DB: Gaps: 7

US-10-019-114A-7 (1-302) x US-08-957-742-2 (1-1660)

QY 60 SerSerValProaspGlyLysProLeuValArgLysProCysArgSerCysAlaValVal 79
DB 510 TCCCTGCTCCCAATGTCGCCCATTAATGAATTAATGTTTGGCTGTGTT 569
QY 80 SerSerSerGlyGlnMetLeuGlySerGlyLeuGlyAlaGluLeuLeuSerAlaGluCys 99

```
Db      570 GGAACAGTGTGATCTTGACAGGAGTGCAGTGTGACAGAAATAGATAATCAGATTTT 629
Qy      100 ValPheArgMetLeuGlnAlaProThrValGlyPheGlnAlaAspValGlyGlnArgSer 119
Db      630 GTTCTCTGATGCAATTTTGCCCGACAGAGCTTTCCACAAAGATGTGGAGAGAAACC 689
Qy      120 ThrLeuArgValValSerHisThrSerValProLeuLeuLeuArgAsnTyrSerHisTyr 139
Db      690 AACCTCACA-----ACCTTCATCCGAGCATCTTAGAGAAATATTACAACAAT 737
Qy      140 PheGlnLysAlaArgAspThrLeuTyrMetValTyrGlyGlnGlyArgHisMetAspArg 159
Db      738 CTTTACCATTTGACACCGTAACAACCTCTCTCCTCAGTTTAAATAAGCTTGATGGGCC 797
Qy      160 ValLeu-----GlyGlyArgThrTyrArgThrLeu 170
Db      798 ATACTTGGATCCCTGCATTTTCTTCCACACTTCTGCACCTGTAACGAGAACCTTAGTG 857
Qy      171 GlnLeuThrArgMetTyrProGly---LeuGlnValTyrThrPheThrGluArgMetLeu 189
Db      858 GATTTTGTGTGAGCACAGAGTCAGTTAAAGTC-----CAGTTG 899
Qy      190 AlaTyrCysAspGlnIlePheGln----AspGluThrGlyAsnArgArgGlnSer 207
Db      900 GCTTGCCCTGGAATATCATGCAACATGTCACACAGTACTGGAATAAACAACCTGTCA 959
Qy      208 GlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeuGluLeuCysGluGlu 227
Db      960 CCCAAGCAGCTGACACAGGATATCTTAATGTATCTTTGATCTGCAATATGTGAAGAG 1019
Qy      228 IleValValTyrGlyMetValSer-----AspSerTyrCysArgGluLysSerHis 244
Db      1020 ATCCACTGTATCGGTTCTGSCCCTTGATTTGACCCCAACACAGGAG----- 1070
Qy      245 ProSerValProTyrHisTyrPheGluLys 254
Db      1071 ---GATCTGCCCTTACCACTACTATGACAAA 1097
```

Search completed: December 21, 2004, 22:52:56
Job time : 113 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 21, 2004, 22:49:40 ; Search time 598 Seconds

(without alignments)

2788.115 Million cell updates/sec

Title: US-10-019-114a-7

Perfect score: 1626

Sequence: 1 MKRGRVLVILICSVFSAV.....FSRMAKRPVFAHPSMRT 302

Scoring table:

BLOSUM62		
Xgapop 10.0 ,	Xgapext 0.5	
Ygapop 10.0 ,	Ygapext 0.5	
Fgapop 6.0 ,	Fgapext 7.0	
Delop 6.0 ,	Delext 7.0	

Searched: 4093002 seqs, 2760418825 residues

Total number of hits satisfying chosen parameters: 8186004

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-Q=/cg2_1/USPTO.spool/US10019114/runat_21122004_100142_3479/app.query.fasta_1.455
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -MATRIX=tblone62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HSPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10019114 @CGN 1 1 480 @runat_21122004_100142_3479
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA:*
1: /cg2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cg2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cg2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cg2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cg2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cg2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cg2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cg2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cg2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cg2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cg2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cg2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq:*
13: /cg2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cg2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cg2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cg2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cg2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
18: /cg2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
19: /cg2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cg2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cg2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1626	100.0	1556	11 US-09-978-360A-2	Sequence 2, Appl1
2	1626	100.0	1556	10 US-10-315-664-25	Sequence 25, Appl
3	1622	99.8	1735	10 US-09-895-298-15	Sequence 15, Appl
4	1605	98.7	1151	16 US-10-262-511-183	Sequence 183, App
5	861	53.0	743	9 US-09-823-245A-30	Sequence 30, Appl
6	812	49.9	966	10 US-09-895-298-121	Sequence 121, App
7	738	45.4	2956	18 US-10-844-874-19	Sequence 19, Appl
8	721	44.3	966	10 US-10-188-186-21	Sequence 21, Appl
9	721	44.3	3283	14 US-10-198-846-11536	Sequence 11536, A
10	720	44.3	3207	15 US-10-104-047-380	Sequence 380, App
11	700.5	43.1	1005	16 US-10-188-186-23	Sequence 23, Appl
12	684	42.1	1106	16 US-10-188-186-19	Sequence 19, Appl
13	639.5	39.3	2048	14 US-10-097-340-188	Sequence 188, App
14	639.5	39.3	2090	15 US-10-094-749-802	Sequence 802, App
15	636.5	39.1	2075	15 US-10-288-252-31	Sequence 31, Appl
16	636.5	39.1	2075	17 US-10-398-038-31	Sequence 31, Appl
17	618	38.0	2017	16 US-10-108-260A-895	Sequence 895, App
18	544.5	33.5	2401	10 US-09-946-374-55	Sequence 55, Appl
19	544.5	33.5	2401	13 US-10-052-586-205	Sequence 205, App
20	544.5	33.5	2401	14 US-10-174-590-205	Sequence 205, App
21	544.5	33.5	2401	14 US-10-176-758-205	Sequence 205, App
22	544.5	33.5	2401	14 US-10-175-737-205	Sequence 205, App
23	544.5	33.5	2401	14 US-10-174-581-205	Sequence 205, App
24	544.5	33.5	2401	14 US-10-176-483-205	Sequence 205, App
25	544.5	33.5	2401	14 US-10-176-749-205	Sequence 205, App
26	544.5	33.5	2401	14 US-10-176-914-205	Sequence 205, App
27	544.5	33.5	2401	14 US-10-176-915-205	Sequence 205, App
28	544.5	33.5	2401	14 US-10-173-706-205	Sequence 205, App
29	544.5	33.5	2401	14 US-10-173-738-205	Sequence 205, App
30	544.5	33.5	2401	14 US-10-175-752-205	Sequence 205, App
31	544.5	33.5	2401	14 US-10-176-482-205	Sequence 205, App
32	544.5	33.5	2401	14 US-10-176-757-205	Sequence 205, App
33	544.5	33.5	2401	14 US-10-174-572-205	Sequence 205, App
34	544.5	33.5	2401	14 US-10-176-913-205	Sequence 205, App
35	544.5	33.5	2401	14 US-10-180-552-205	Sequence 205, App
36	544.5	33.5	2401	14 US-10-180-557-205	Sequence 205, App
37	544.5	33.5	2401	14 US-10-173-700-205	Sequence 205, App
38	544.5	33.5	2401	14 US-10-174-579-205	Sequence 205, App
39	544.5	33.5	2401	14 US-10-174-582-205	Sequence 205, App
40	544.5	33.5	2401	14 US-10-174-588-205	Sequence 205, App
41	544.5	33.5	2401	14 US-10-175-739-205	Sequence 205, App
42	544.5	33.5	2401	14 US-10-175-740-205	Sequence 205, App
43	544.5	33.5	2401	14 US-10-175-743-205	Sequence 205, App
44	544.5	33.5	2401	14 US-10-176-488-205	Sequence 205, App
45	544.5	33.5	2401	14 US-10-176-492-205	Sequence 205, App

ALIGNMENTS

RESULT 1
US-09-978-360A-2
; Sequence 2, Application US/09978360A
; Publication No. US20040110939A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Uobert, Severin
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56 US4.CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563

LOCATION: 1524..1556
US-10-315-664-25

Alignment Scores:	
Pred. No.:	1.17e-196
Length:	1556

US-10-019-114A-7 (1-302) X US-10-315-664-25 (1-1556)

QY	1	MetysalAProGlyArgLeuValLeuIleIleLeuCySerValValPheSerLAlaVal	20
Db	261	ATGAAGGCTCGGGGCTGGGCTGCTGCTACTAATCTGTCTCCGTGCTTCTCTGCGCTC	320
QY	21	TyrIleLeuLeuCyCyValTPrLaglLeuProLeuCyLeuAlaThrCyLeuAlaPheI	40
Db	321	TACATCTCTCTGTGTGTGTGTGGGCCCGGCTCGCCCTCTGTGGACACGTGTCGACAC	380
QY	41	HisPheProThrArgIysPheArgProThrValProGlyProLeuHisPheSerGlyIysSer	60
Db	381	CACCTTCCCAAGGCTCCAGGCCACTGTGGCCGGAACCTGTCACTTCACTGTAATTAAGC	440
QY	61	SerValProAspGlyIysProLeuValArgIuProCyArgSerCyenAlaValAlaSer	80
Db	441	AGTGGCCAGATGGGAAGCCGCTGGATCGCGAGCCCTGCCGACGTGAGCCGTGTATGTC	500
QY	81	SerSerGlyGlnMetLeuGlySerGlyLeuGlyAlaGluIleAspSerLAlaGlyVal	100
Db	501	AGCTTCGGCCAAATGCTGGGGCTCAGGCTGTGGGTGTGAGATCGAAGTCCCAAGTCCGG	560
QY	101	PheArgMetAngIAlaProThrValGlyPheGluAlaAspValGlyGlnArgSerThr	120
Db	561	TTCCGCATGAACAGGCCCCCAACCGTGGCTTTAGAGCGATGTGGCCAGGCAAGCAC	620
QY	121	LeuArgValValSerHisThrSerValProLeuLeuLeuArgAsnTyrSerHisTyrPhe	140
Db	621	CTGGGTCTGCTCTACACACAAAGCTGGCTGCTGTCGCGCACTAATTCACACTACCTC	680
QY	141	GlnIysAlaIaArgAspThrLeuTyrMetValTPrGlyGlnGlyArgHisMetAspArgVal	160
Db	681	CAGAAGCCCGAGACACGCTCTACATGATGTGGGCCAGAGGCCAGCACTGGACCGGGT	740
QY	161	LeuGlyGlyArgThrTyrArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGln	180
Db	741	CTCGCGGCGCAGCACTACCGCAGCGCTGTGAGCTCACCAAGATGTACCCGGGCTGCAG	800
QY	181	ValTyrThrPheThrArgIuArgMetMetLAlaTyrCyAspArgIlePheGlnAspGluThr	200
Db	801	GTTGACACCTTCAAGAGCGCATATGATGCTTACGCAACAGATTTCCAGACCAAGAG	860
QY	201	GlyIysAsnArgArgGlnSerGlySerPheLeuSerThrGlyTyrPheThrMetIleLeu	220
Db	861	GGCAAGAACGGGAAGCAGTGGGCTCTTCTCTCAGCAACCGGCTGTTCACATGATCTC	920
QY	221	AlaLeuGlnLeuCyGluGlnIleValValTyrGlyMetValSerAspSerTyrCyArg	240
Db	921	GCGCTGAGCGTGTGAGAGATGTTGTTATGAGTGAAGTCAAGCAAGCTACGTCAGAG	980
QY	241	GluIysSerHisPheProSerValProTyrHisTyrPheGlyIuValGlyArgLeuAspGluCys	260
Db	981	GAGAAAGCAACCCCTCAATGCTTTCACACTTATGAAAGGGCCGGGCTGATATAGTGT	1040
QY	261	GlnMetTyrLeuAlaHisGluGlnAlaProArgSerAlaHisArgPheIleThrGluIys	280
Db	1041	CAGATGTACCTGGCACAAGACAGGCGCCGGAAGCGCCACCGCTTATCACTAGAAG	1100
QY	281	AlaValPheSerArgTPrLAlaIuValuArgProIleValPheAlaHisProSerTPrArg	300
Db	1101	GCGGTCTTCTCCGCTGGCCAAAGAGAGGCCCATGTGTTCGCCATCGCTCTGGAGG	1160

QY	301	ThrGlu	302
Db	1161	ACTGAG	1166

RESULT 3

US-09-895-298-15
; Sequence 15, Application US/09895298

Publication No. US20030078405A1

; GENERAL INFORMATION:

APPLICANT: ROSEN, et al.
TITLE OF INVENTION: 47

FILE REFERENCE: PZ035P1

CURRENT APPLICATION NUMBER: US/09/895,298

CURRENT FILING DATE: 2001-07-02

;; PRIOR APPLICATION NUMBER: 09/591,16
PRIOR FILING DATE: 2000-06-09

PRIOR APPLICATION NUMBER: PCT/US99/29950

PRIOR FILING DATE: 1999-12-16

;
PRIOR APPLICATION NUMBER: 60/113,006

PRIOR FILING DATE: 1998-12-18

;; PRIOR APPLICATION NUMBER: 60/112,809
PRIOR FILING DATE: 1999-12-17

NUMBER OF SEO ID NOS: 231

```

; SOFTWARE: PatentIn Ver. 2.0

```

; SEQ ID NO 15

LENGTH: 1735
TYPE: DWT

```

;
; TYPE: DNA
; ORGANISM: Homo sapiens

```

FEATURE:

NAME/KEY: SITE

LOCATION: (1002)

OTHER INFORMATION: n equals a,t,g, or c
HS-09-895-298-15

C
T
O
C
U
O
C
V
C
C

Alignment Scores:

Pred. No.:	4.45e-196	Length
------------	-----------	--------

Score:	1622.00	Matches:	301
Percent Similarity:	99.67%	Conservative:	0

US-10-019-114A-7 (1-302) X US-09-895-298-15 (1-1735)

QY	1	MeLrVSAIAProGIYArlgUeuValleuIlelleuCYsserValValPheSerAVal	20
Db	293	ATGAAGGCTCCGGAGCGGCTCGTGTCAATCTGTGGTCCGGCTTCTTCGCGGTC	352
QY	21	TYrIleLeuLeuCYsCTrAlaIglYeuProleuCYsleuAlaTrnCYSLeuAphIs	40
Db	353	TACATCTCTGTGTGCTGGCGGCGCTGCGCCCTGTGGCTGGCCACTGCTGGACAC	412
QY	41	HisPheProThGISeSerArProThrValProGIYProleuHisPheSerGIYTrSer	60
Db	413	CACCTTCCCAAGGCTCCAGGCCACTGTGGCGGAGCCCTGCACCTTCAAGTAAATAGC	472
QY	61	SeValProAspGIYLePProleuValArgIuProCYsArGserCYsAlaValISer	80
Db	473	AGTGTCCAGATGGAGAGCGGTGTCCGCAAGCCTTGGCAGACTGTGCGGTGTCC	522
QY	81	SeSerGIYGIuMetLeuGIYSerGIYLeuGIYAlaIuIleAsPSeValAGIuCYsVal	100
Db	533	AGCTCCGGCAAAATCTGGGCTCAGGCGCTGGGTCTAGATCGACATGCGAGTGGTG	592
QY	101	PheArGMeAsnGIuAlaProThrValGIYpHeGIuAlaAsPValGIYGIuArGserThr	120
Db	593	TTCCGGATGAACCAAGGCCCCACCGTGGCTTTAAGGCGATGTGGGCCAGCGGAGACC	652
QY	121	LeuArGValValSerHisThrSerValProleuLeuLeuArgAntYSerHisTrYPhe	140
Db	653	CTGGGTGTGTCTCACACAAAGCGTGGCTGTCTGGCAACTATTCACACTACTTC	712
QY	141	GIuLYeAlaArgAsPThrLeuTYrMeValTrpGIYGIuGIYArGHisPheAsPArgVal	160

```

Db      713 CAGAAAGCCCGAAGACAGCTCTACATGCTGGAGGCGAGGACATGAGCCGGGTG 772
Qy      161 IeuglYglYargThrYrargThrLeuGlIleuThzrArgMetYrProglYleuGln 180
Db      773 CTGGGGCGGCGGACCTACCGGACGCTCTGAGTCCACGAGATGATCCCGGCGCTGAG 832
Qy      181 ValYrThrPheThrGluArgMetMetAlaYrCyAspGlnIlePheGlnAspGluThr 200
Db      833 GTGTACACCTTCAAGGAGCCATGATGGCTTACGACACCAAGATCTTCCAGGAGCAGACG 892
Qy      201 GlyYAspArgArgArgGlnSerGlySerPheLeuSerThrGlyThrPheThrMetIleLeu 220
Db      893 GGCAGAGACCGGAGGAGCGTGGGCTCTTCTTCAAGACCGGCTGTTCACCATGATCTTC 952
Qy      221 AlaLeuGluLeuGlyGlnGluIleValYrGlyMetValSerAspSerYrCyAspArg 240
Db      953 GCGCTGAGAGCTGTGTAGAGATCGTGTCTATGGAGATGGTCAGCGACATCTACTGAGG 1012
Qy      241 GluYsSerHisProSerValProYrHisYrPheGluYsglyYargLeuAspGluCys 260
Db      1013 GAGAGAGCCACCCCTCAGTGGCTTACCTACTTGTGAGAGAGCGCGCTAGATGATGT 1072
Qy      261 GluMetYrLeuAlaHisGluGlnAlaProArgSerAlaHisArgPheIleThrGluYs 280
Db      1073 CAGATGATCTGGACGACGAGCGAGCGCCCGAGAGCCGACCGCTTCACTCATCTGAGAG 1132
Qy      281 AlaValPheSerArgTrrAlaYsYsArgProIleValPheAlaHisProSerTrrArg 300
Db      1133 GCGGCTTCTTCCCGCTGGGCGAAGAGAGGCCATCGTGTTCGCCATCTCCTGAGAG 1192
Qy      301 ThrGlu 302
Db      1193 ACTGAG 1198

```

RESULT 4

US-10-262-511-183
 Sequence 183, Application US/10262511
 Publication No. US20040038222A1

GENERAL INFORMATION:

APPLICANT: Smithson, Glendia
 APPLICANT: Millet, Isabelle
 APPLICANT: Peyman, John A.
 APPLICANT: Kekuda, Ramesh
 APPLICANT: Ju, Jingtang
 APPLICANT: Li, Li
 APPLICANT: Guo, Xiaojia (Sasha)
 APPLICANT: Patnirajan, Meera
 APPLICANT: Spytek, Kimberly A.
 APPLICANT: Edinger, Shlomit R.
 APPLICANT: Ellerman, Karen
 APPLICANT: Malyankar, Uriel M.
 APPLICANT: Ort, Tatiana
 APPLICANT: Gorman, Linda
 APPLICANT: Zerhusen, Bryan D.
 APPLICANT: Anderson, David W.
 APPLICANT: Zhong, Mei
 APPLICANT: Catterton, Elina
 APPLICANT: Ji, Weizhen
 APPLICANT: Miller, Charles E.
 APPLICANT: Raetelli, Luca
 APPLICANT: Stone, David J.
 APPLICANT: Pena, Carol E. A.
 APPLICANT: Shenoy, Suresh G.
 APPLICANT: Shimkets, Richard A.
 APPLICANT: Rothenberg, Mark E.
 APPLICANT: Leach, Martin D.
 APPLICANT: Agee, Michele L.
 APPLICANT: Beghe, Constance
 TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 FILE REFERENCE: 21402-462C
 CURRENT APPLICATION NUMBER: US/10/262, 511
 CURRENT FILING DATE: 2003-05-28

```

PRIORITY APPLICATION NUMBER: 60/326,483
PRIORITY FILING DATE: 2001-10-02
PRIORITY APPLICATION NUMBER: 60/373,815
PRIORITY FILING DATE: 2002-04-19
PRIORITY APPLICATION NUMBER: 60/327,917
PRIORITY FILING DATE: 2001-10-09
PRIORITY APPLICATION NUMBER: 60/381,642
PRIORITY FILING DATE: 2002-05-17
PRIORITY APPLICATION NUMBER: 60/328,029
PRIORITY FILING DATE: 2002-10-09
PRIORITY APPLICATION NUMBER: 60/381,038
PRIORITY FILING DATE: 2002-05-16
PRIORITY APPLICATION NUMBER: 60/328,056
PRIORITY FILING DATE: 2001-10-09
PRIORITY APPLICATION NUMBER: 60/373,260
PRIORITY FILING DATE: 2002-04-17
PRIORITY APPLICATION NUMBER: 60/373,826
PRIORITY FILING DATE: 2002-04-19
PRIORITY APPLICATION NUMBER: 60/327,435
PRIORITY FILING DATE: 2001-10-05
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 439
SOFTWARE: CuroSeqList version 0.1
SEQ ID NO 183
LENGTH: 1151
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (421)..(1074)
FEATURE:
NAME/KEY: misc feature
LOCATION: (2)..(2)
OTHER INFORMATION: wherein n may be a, c, g or t
OTHER INFORMATION: wherein n may be a, c, g or t
US-10-262-511-183

```

Alignment Scores:

Pred. No.: 3,58e-194 Length: 1151
 Score: 1605.00 Matches: 298
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.71% Indels: 0
 DB: Gaps: 0

US-10-019-114A-7 (1-302) x US-10-262-511-183 (1-1151)

```

Qy      5 GlyArgLeuValIleuIleLeuCysSerValValPheSerAlaValTyrIleLeu 24
Db      181 GGTGGGCTGTGTCTCATCATCTGTGCTCGTGTCTTCTTCCGCTTACATCTCTG 240
Qy      25 CysCysTrrAlaGlyLeuProLeuCysLeuAlaThrCysLeuAspHisPheProThr 44
Db      241 TGCTGTGGGCGGCGGCTGCTCTGCTGGCAGCAGCTGTGACACCACTTCCGACACA 300
Qy      45 GlySerArgProThrValProGlyProLeuHisPheSerGlyYrSerSerValProAsp 64
Db      301 GGCCTCCAGGCCCATGTCGCGGAGCCCTGCTCATCTTCACTGATGATATGACATGTGCGAGAT 360
Qy      65 GlyYsPheProLeuValArgGluProCysArgSerCysAlaValAlaSerSerGlyGln 84
Db      361 GGGAGAGCGGTGTGTGCGGAGCCCTGCGGAGCTGTGCTGTGTGTCCAGCTCCGGCAA 420
Qy      85 MetLeuGlySerGlyLeuGlyAlaGluIleAspSerAlaGluCysValPheArgMetAsn 104
Db      421 ATGCTGGGCTCAGAGCTGGTGTCTGAGATGACAGTCCGAGTGGCGGTCCGATTAAC 480
Qy      105 GlnAlaProThrValGlyPheGluAlaAspValGlyGlnArgSerThrLeuArgValVal 124
Db      481 CAGGCGGCCACCGTGGGCTTTGAGGCGGATGTGGGCGAGCGACACCTGCGTGTCTC 540

```

QY 125 SerHisSerValProLeuLeuLeuArgSerHisTyrPheGlnValAlaArg 144
 Db 541 TCACACCAAGCGCGCTCTGCTGCAACTATTCACTACTCCAGAAAGCCCGA 600
 QY 145 AspThrLeuTyrMetValTrrpGlnGlnYarGHisMetAspArgValAlaGlnYarG 164
 Db 601 GACACGCTTACAGTGTGTGGGGCCAGGGCAGACATGACCGGGTGTGGGGCCGC 660
 QY 165 ThrTyrArgThrLeuGlnLeuThrArgMetTyrProGlnLeuGlnValTyrThrPhe 184
 Db 661 ACCAACCGCAGCTGTGAGACTCACCAGATGTACCCTGGCTGTGAGTGTACACCTTC 720
 QY 185 ThrGlnArgMetMetAlaTyrCyAspGlnIlePheGlnAspGlnTrrGlnYarGAsnArg 204
 Db 721 ACCGAGGCGATGAGGGCTTCTGCGACAGATCTTCCAGACGAGACGGGCGCAACCGG 780
 QY 205 ArgGlnSerGlySerPheLeuSerThrGlyTrrPheThrMetIleLeuAlaLeuGlnLeu 224
 Db 781 AGCGAGTCGGGCTCTCTCTCAGACCGGCTGGTTCCATGATCTTGGCTGGAGCTG 840
 QY 225 CysGlnGlnIleValValTrrGlyMetValSerAspSerTyrCyAspGlnYarGHis 244
 Db 841 TGTGAGGAGATCGTGTGTGTGATGTGTGACGACAGCTACTGACGAGGAGAGAGCCAC 900
 QY 245 ProSerValProTyrHisTrrPheGlnYarGLeuAspGlnYarGlnMetTrrLeu 264
 Db 901 CCCCTACATGCTTACCACTACTTGTGAGAGGGCCGGCTAGATGTGATGATGTACTTC 960
 QY 265 AlHisGlnGlnIleProArgSerAlaHisArgPheIleThrGlnYarGAlaPheSer 284
 Db 961 GCACACAGCAGGCGCCCGAGCGCCACCGCTTCACTGAGAAAGCGGCTTCTCC 1020
 QY 285 ArgTrrAlaYarYarGProIleValPheAlaHisProSerTrrArgThrGln 302
 Db 1021 CGCTGGGCGCAAGAGAGGCCCATGTGTTCGCCCATCTGCTGAGAGACTGAG 1074

RESULT 5
 US-09-823-245A-30
 ; Sequence 30, Application US/09823245A
 ; Publication No. US20020039760A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wong, Gordon G.
 ; APPLICANT: Clark, Hilary
 ; APPLICANT: Fechtel, Kim
 ; APPLICANT: Agostino, Michael J.
 ; APPLICANT: Howes, Steven H.
 ; APPLICANT: Resnick, Richard J.
 ; APPLICANT: Gulukota, Kamalakkar
 ; APPLICANT: Graham, James R.
 ; APPLICANT: Genetics Institute, Inc.
 ; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
 ; FILE REFERENCE: GIN 6401
 ; CURRENT APPLICATION NUMBER: US/09/823,245A
 ; PRIOR FILING DATE: 2001-03-29
 ; PRIOR APPLICATION NUMBER: 60/194,941
 ; NUMBER OF SEQ ID NOS: 631
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 30
 ; LENGTH: 743
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-823-245A-30

Alignment Scores:
 Pred. No.: 1,61e-99 Length: 743
 Score: 861.00 Matches: 161
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 52.95% Indels: 0
 DB: 9 Gaps: 0

US-10-019-114A-7 (1-302) x US-09-823-245A-30 (1-743)
 QY 1 MetIysAlaProGlnYarGLeuValIleIleLeuCySerValAlaPheSerAlaVal 20
 Db 260 ATGAAAGGCTCCGGGTGGCTGTGTCTCATCATCTGTGTCCGGTCTTCTCTGGCTC 319
 QY 21 TyrIleLeuLeuCySerTrrPalaGlnYarLeuProLeuCyLeuAlaTrrCyLeuAspHis 40
 Db 320 TACATCTCTCTGT 379
 QY 41 HisPheProThrGlySerArgProThrValProGlyProLeuHisPheSerGlyTrrSer 60
 Db 380 CACTTCCACAGGCTTCAAGGCCCATCTGTGCGGAGCCCTGCACTTCAATGATATAC 439
 QY 61 SerValProAspGlnYarProLeuValArgGlnProCyAspSerGlyAlaValAlaSer 80
 Db 440 AGTGTCCAGATGGAGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 499
 QY 81 SerSerGlyGlnMetLeuGlnSerGlyLeuGlnYarGlnIleAspSerAlaGlnYarVal 100
 Db 500 AGCTCCGGCCAAATGTGGGCTCAGGCTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 559
 QY 101 PheArgMetAsnGlnIleProThrValGlyPheGlnIleAspValGlyGlnArgSerThr 120
 Db 560 TTCCGATGAAACAGGCGCCACCGTGGCTTTGAGCGGATGTGGCCAGCGACACACC 619
 QY 121 LeuArgValAlaSerHisTrrSerValProLeuLeuLeuArgSerHisTrrPhe 140
 Db 620 CTGGT 679
 QY 141 GlnYarAlaArgAspThrLeuTrrMetValTrrpGlnGlnYarGHisMetAspArgVal 160
 Db 680 CAGAAAGCCCGAAGACACACTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 739
 QY 161 Leu 161
 Db 740 CTC 742

RESULT 6
 US-09-895-298-121
 ; Sequence 121, Application US/09895298
 ; Publication No. US20030078405A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: 47 Human Secreted Proteins
 ; FILE REFERENCE: P2035P1
 ; CURRENT APPLICATION NUMBER: US/09/895,298
 ; PRIOR FILING DATE: 2001-07-02
 ; PRIOR APPLICATION NUMBER: 09/591,16
 ; PRIOR FILING DATE: 2000-06-09
 ; PRIOR APPLICATION NUMBER: PCT/US99/29950
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: 60/113,006
 ; PRIOR FILING DATE: 1998-12-18
 ; PRIOR APPLICATION NUMBER: 60/112,809
 ; PRIOR FILING DATE: 1998-12-17
 ; NUMBER OF SEQ ID NOS: 231
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 121
 ; LENGTH: 966
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-895-298-121

Alignment Scores:
 Pred. No.: 4,25e-93 Length: 966
 Score: 812.00 Matches: 153
 Percent Similarity: 99.35% Conservative: 0
 Best Local Similarity: 99.35% Mismatches: 1
 Query Match: 49.94% Indels: 1
 DB: 10 Gaps: 0

US-10-019-114A-7 (1-302) x US-09-895-298-121 (1-966)

QY 149 MetValTrpGlyGlnGlyArgHisMetAspArgValLeuGlyValArgThrTyrArgThr 168
 Db 3 ATGGTGGGGGCGAGGGGCGAGCAGATGACCGGGTGTCTCGCGCGCGACCTTACCGGACG 62
 QY 169 LeuLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyrThrPheThrGluArgMet 188
 Db 63 CTGCTGAGCTCACAGAGATGATCCCGGCTGAGGTGTACACTTCAAGAGCGCATG 122
 QY 189 MetAlaTyrCysAspGlnIlePheGlnAspGluThrGlyLysAsnArgArgGlnSerGly 208
 Db 123 ATGGCTACTGCGACAGATCTTCCAGAGCAGAGCGGCAAGAACCGGAGCGAGTCCGGGC 182
 QY 209 SerPheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeuGlyLeuGlyGlnIle 228
 Db 183 TCTCTTCCAGACCGCGCTGGTTCACCATGATCTCCGCTGAGCTGTGTAGAGAGATC 242
 QY 229 ValValTyrGlyMetValSerAspSerTyrCysArgGlyLysSerHisProSerValPro 248
 Db 243 GTGGTCTATGGAGTGTACAGCAGAC-TACTGCGAGGAGAGAGCCACCTTCACTGCT 301
 QY 249 TyrHisTyrPheGlyLysGlyArgLeuAspGluCysGlnMetTyrLeuAlaHisGlyGln 268
 Db 302 TACCACTACTTTGAGAGGGCGGCTAGATGATGTACAGATGTACTGCGACAGCAGCAG 361
 QY 269 AlaProArgSerAlaHisArgPheIleThrGlyLysAlaValPheSerArgTTPAlaLys 288
 Db 362 GCGGCCGGAAGCGGCCACCGCTTCATCATGAGAGGCGGCTTCCCGCTGGGCCAAG 421
 QY 289 LysArgProIleValPheAlaHisProSerTyrPheThrGlu 302
 Db 422 AAGAGGCCCATCGTGTTCGCCCATCCCTCTGAGAGACTGAG 463

RESULT 7
 US-10-844-874-19
 ; Sequence 19, Application US/10844874
 ; Publication No. US20040204381A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Moskal, Joseph
 ; APPLICANT: Yamamoto, Hirokazu
 ; TITLE OF INVENTION: Detection and Treatment of Glyco-Enzyme-Related Disease
 ; FILE REFERENCE: 97-186-E
 ; CURRENT APPLICATION NUMBER: US/10/844,874
 ; CURRENT FILING DATE: 2004-05-13
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 19
 ; LENGTH: 2956
 ; TYPE: DNA
 ; ORGANISM: Rat
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: rat alpha 2, 6-ST
 US-10-844-874-19

Alignment Scores:
 Pred. No.: 6,16e-83 Length: 2956
 Score: 738.00 Matches: 147
 Percent Similarity: 69.09% Conservative: 43
 Best Local Similarity: 53.45% Mismatches: 77
 Query Match: 45.39% Indels: 8
 DB: 18 Gaps: 5

US-10-019-114A-7 (1-302) x US-10-844-874-19 (1-2956)
 QY 30 LeuProLeuCysLeuAlaThrCysLeuAspHisPheProThrGlySerArg---Pro 48
 Db 320 CTTTCCCTTGTCTCTGAACTGCTTTGGACAACCAATGATCCCTTTGTCTT 379
 QY 49 ThrValProGly--ProLeu-----HisPheSerGlyTyrSerSerValProAspGlyL 66
 Db 380 ACACACTCAGGAGGAGCTCTTCAAACTCACTAT---GGATACATATAAAGTGAAGACCAAG 436

QY 66 YsrProLeuValArgGluProCysArgSerCysAlaValIleSerSerGlyGlnMetL 86
 Db 437 AGCCTTTG---CAGCTGAAGTGTAACTTGTCCCTGTGTGAACTGGGTCAATGG 493
 QY 86 euGlySerGlyLeuGlyAlaGluIleAspSerAlaGluCysValPheArgMetAsnGlnA 106
 Db 494 TTGGACAGAAAGGTGGGGAAGAGATAGACCGTCCCTGCACTGAGAAATGAACATG 553
 QY 106 LprOThrValGlyPheGlnAlaAspValGlyGlnArgSerThrLeuArgValIleSerH 126
 Db 554 CCCCACCAAGGTTTTCAGAAAGATGTGGCTACATGACATGATGTCCTCCAGTGTCTAC 613
 QY 126 LstHrSerValProLeuLeuArgAsnTyrSerHisTyrPheGlnLysAlaArgAspT 146
 Db 614 ACACGAGCGTCCCTCTTGTCTGAADACCCGACATATTTTTCAGAGAGCGACACAA 673
 QY 146 hrLeuTyrMetValTyrGlyGlnGlyArgHisMetAspArgValLeuGlyValArgThr 166
 Db 674 CCATTACGTCATCTGGGCGCTTTCGCAACATGAGAGAGAGCGGAAAGGATCGGT 723
 QY 166 YrArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyrThrPheThrG 186
 Db 734 ACAACATGTCTAAAGAGACGCTGACGCTTACCCAGACGCGCATCTACATGACACAG 723
 QY 186 LurArgMetAlaTyrCysAspGlnIlePheGlnAspGluThrGlyLysAsnArgArg 206
 Db 794 AGCAGCGAGTGAAGTACTGAGAGGGGTGTTTAAGATGAATACTGGGAAGACAGAGTCC 853
 QY 206 InSerGlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeuGlyLeuGly 226
 Db 854 AGTGGCTCTCTATCTACAGCAGCGGCTTCACTTATCCCTGGCCATGAGAGCCCTGCT 913
 QY 226 IuGluIleValTyrGlyMetValSerAspSerTyrCysArgGlyLysSerHisProS 246
 Db 914 ACAATATCCAGCTACGGGATATATGAGCTTACTGACAGACAGAAAGGTATAGAA 973
 QY 246 eValProTyrHisTyrPheGlyLysGlyArgLeuAspGluCysGlnMetTyrLeuAlaH 266
 Db 974 AAGTCCCTTCTACTTACTATGAACAGAGGAG---GACGAGTGTAAAGATCTTCTCC 1030
 QY 266 IsgGluGlnAlaProArgSerAlaHisArgPheIleThrGlyLysAlaValPheSerArgT 286
 Db 1031 ATGAACATGCCCATTCGCGGAGACACGGTTCATCACCGGAGAAAGTGTGGCAAGT 1090
 QY 286 rPalAlaLysArgProIleValPheAlaHisProSerTrp 299
 Db 1091 GGGCCAGAAACACAGATATGTATTAACACACCCCAACTGG 1131

RESULT 8
 US-10-188-186-21
 ; Sequence 21, Application US/10188186
 ; Publication No. US20040029789A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson et al.
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 ; FILE REFERENCE: 21402-397C
 ; CURRENT APPLICATION NUMBER: US/10/188,186
 ; CURRENT FILING DATE: 2002-07-02
 ; PRIOR APPLICATION NUMBER: 60/303046
 ; PRIOR FILING DATE: 2001-07-05
 ; PRIOR APPLICATION NUMBER: 60/360814
 ; PRIOR FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: 60/303828
 ; PRIOR FILING DATE: 2001-09-07
 ; PRIOR APPLICATION NUMBER: 60/323380
 ; PRIOR FILING DATE: 2001-09-19
 ; PRIOR APPLICATION NUMBER: 60/361133
 ; PRIOR FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: 60/304016
 ; PRIOR FILING DATE: 2001-07-09
 ; PRIOR APPLICATION NUMBER: 60/304502
 ; PRIOR FILING DATE: 2001-07-11
 ; PRIOR APPLICATION NUMBER: 60/305262

```

: PRIOR FILING DATE: 2001-07-13
: PRIOR APPLICATION NUMBER: 60/373981
: PRIOR FILING DATE: 2002-04-19
: PRIOR APPLICATION NUMBER: 60/305673
: PRIOR FILING DATE: 2001-07-16
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 368
: SOFTWARE: Custom
: SEQ ID NO: 21
: LENGTH: 966
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (19)..(934)
US-10-188-186-21

Alignment Scores:
Pred. No.: 1,74e-81 Length: 966
Score: 721.00 Matches: 139
Percent Similarity: 65.41% Conservative: 52
Best Local Similarity: 47.60% Mismatches: 73
Query Match: 44.34% Indels: 28
DB: 16 Gaps: 6

US-10-019-114A-7 (1-302) x US-10-188-186-21 (1-966)

QY 26 CysTpaIaGlyLeuProLeuCySleuAla----- 35
Db 83 TGCTGG-----TTGTGCGCTTGTAATGAAGTAATTCCTGCTACTAA 130

QY 36 ThrCySleuAap-----HisHsPheProThrlGlySerAapProThr 49
Db 131 ACTGCTTTGGAAACCTGGTACAAAGTGAGTACCATCTCTCAACATCA----- 181

QY 50 ValPProGlyProLeuHisPheSer--GlyTySerSerValProAapGlyLeuProLeu 69
Db 182 ---GGCGGCCCTTGGAACTCACTAGTGAATCAATAATGTGAAGACACAAAGCCTTGC 238

QY 69 aIArgGlyProCyArGserCySAlaValValSerSerSerGlyGlnMetLeuGlySerg 89
Db 239 AAGTGCAC---TGTGACCTTTGTGCATAGTGTCAAACTCAGGTCAAGTGTGGCCAGA 295

QY 89 lYleuGlyAlaGlnLeuAapSerAlaGlnCySAlaPheArGhMetAanglnAlaProThrv 109
Db 296 AGGTGGGAATGAGATAGATCGATCGCTCTCGCATTTGGAGAAATGAACAATGCCACCA 355

QY 109 aIglyPheGlnAlaAerValGlyGlnArGserThrlLeuArgValValSerHisThSery 129
Db 356 AAGGTTATGAAGAAATATGCGGCCGATGACCAATGTTGAGTTGTGCCATACCAAGCG 415

QY 129 aIProLeuLeuLeuArGAntYrSerHisTyPheGlnIlySAlaAaPArThrlLeuTyM 149
Db 416 TTCCTCTTTGCTAAAAAACCTGATATATTTTTCAGAGAAGCAATACCTATATTAG 475

QY 149 etValTTPGlyGlnGlyArGHisMetAerAaValLeuGlyGlyArGThrvYrArGThrl 169
Db 476 TTATTTGGGAGCCTTCCGCAATATGAGGAAGAATGCAATGGCATTGTTACAACATG 535

QY 169 eulLeuGlnLeuThrvrGmetYrProGlyLeuGlnValYrThrvPheThrlGlnArGMeCh 189
Db 536 TGAATAAAACAGTGGTATCTATCCGAATGCCCAATATATACGTGACACAGAAACCGCA 595

QY 189 eCaIATyCySArPbInIlePheGlnAaPbGlnThrlGlySAlaAaAaGArGlnSerGlyS 209
Db 596 TGAATTAATGATGATGAGATTTTAAAGAAAGAACTGGAGAAGACAGAGTCAGTGGCT 655

QY 209 erPheLeuSerThrlYrPheThrlMetCileuAlaLeuGlnLeuCySglnGlnIleY 229
Db 656 CATATCTCAGACACAGGTGGTATTACCTTCTCGGCATGACGCGCTGTATGCGATTTC 715

QY 229 aIValITyGlyMetValSerAapSerYrCyArGlnIlySerHisPProSerValProt 249

```

```

Db      716  ACGTTACGGGATGATTAATATGACACCTACTGTAAGAACAAGGGTATPAGAAAAGTCCCT
QY      249  YTHSYTRPheGluLyseGlyValGLeuAspGluCysGlnMetYrLeuAlaHisGluGlnA
Db      776  ACCATTATTAATGAAACAGAGMAA---GATGAGCTGTGATGAATAATTTCTTCATGAACATG
QY      269  lAProArSerAlaHisArpPheIleThnGluLyAlaValPheSerArGTrpAlaLySt
Db      833  CCCCATTAGGGGGTCAATAGGTTTATTCACCTGAAAGAAAGATGTTTGTCTAATAGGGCCACAGA
QY      289  ySArGProIleValPheAlaHisProSerTrp 299
Db      893  AGCAACAGATTAATATTTCACATCCAAACTGG 924

RESULT 9
US-10-198-846-11536
; Sequence 11536, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11536
; LENGTH: 3283
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 3283
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11536

Alignment Scores:
Pred. No.: 1.06e-80 Length: 3283
Score: 721.00 Matches: 139
Percent Similarity: 65.41% Conservative: 52
Best Local Similarity: 47.60% Mismatches: 73
Query Match: 44.34% Indels: 28
DB: 14 Gaps: 6

US-10-019-114A-7 (1-302) x US-10-198-846-11536 (1-3283)

QY      26  CysTrpAlaGlyLeuProLeuCysLeuAla----- 35
Db      184  TGCTGG-----TTGGCGCTTGTGTAATAAGATGAATTTCCATGCTACTAA 231
QY      36  ThrCysLeuAsp-----HisHspheProThnGlySerArGProThn 49
Db      232  ACTGCTTTGGACAACTGCTGACAAAGTGATACCAATTCCTCTACACATACA----- 282
QY      50  ValProGlyProLeuHisPheSer--GlyYrSerSerValProAspGlyLyProLeuV 69
Db      283  ---GGCGGCCCCCTCGAAGCTACCTATGATACATATAATGTGAAGACACAAGAGCCTTTGC 339
QY      69  AlaGgGluProCysArGserCysAlaValValSerSerSerGlyGlnMetLeuGlySerG 89
Db      340  AACTGAGC---TGTCACCTTTGTGTCATAGTCTCAACACAGGTCAAGATGGTTGGCCAGA 396
QY      89  LyLeuGlyAlaGluIleAspSerAlaGluCysValPheArGMetArEnGlnAlaProThnV 109
Db      397  AGGTGGGAATGAGATAGATGATGATCTCTCTGATTTGGAGATGAACAAATGCCCCACCA 456

```



```

PRIOR APPLICATION NUMBER: 60/323380
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: 60/361133
PRIOR FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: 60/304016
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/304502
PRIOR FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: 60/305262
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: 60/373881
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/305673
PRIOR FILING DATE: 2001-07-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 368
SOFTWARE: Custom
SEQ ID NO 23
LENGTH: 1005
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (19)..(973)
US-10-188-186-23

Alignment Scores:
Pred. No.: 7,6e-79
Score: 700.50
Percent Similarity: 62.30%
Best Local Similarity: 43.08%
Query Match: 16
DB: 16

US-10-019-114a-7 (1-302) x US-10-188-186-23 (1-1005)

QY 26 CysTPrAlaGlyLeuProLeuCybLeuAla----- 35
Db 83 TGCCTGG-----TTGTGCGTCTGTAAATGAAGGAATTCCCTGCTACTAA 130

QY 36 ThCysLeuAsp-----HisHisPheProThrGlySerArgProThr 49
Db 131 ACTGCTTTGGACAACCTGACCAAGTGAATACCATCTCTCTACACATACA----- 181

QY 50 ValProGlyProLeuHisPheSer-GlyTyrSerSerValProAspGlyLeuProLeu 69
Db 182 ---GGGGGGCCCTTCGAACTCACTATGATATCAATTAATGTAAGACACAAAGAGCTTTGC 238

QY 69 alArgGlyProCybArgSerCybAlaValAlaSerSerSerGlyGlnMetLeuGlySerG 89
Db 239 AACTGAC--TGTGACCTTTGTGCCATGAGTCAAACTGACAGTCAAGATGGTGGCCAGA 295

QY 89 lylLeuGlyAlaGlyLeuAspSerAlaGlyCysValPheArgMetAsnGlnAlaProThr 109
Db 296 AGCTGGGAATATGATGATGATGATCCCTGCTGATTTGGAAATGAACAATGCCCCACCA 355

QY 109 alGlyPheGlyAlaAspValGlyGlnArgSerThrLeuArgValAlaSerHisThrSer 129
Db 356 AAGGTTATGAAGAAGATGTGGCGCCGATGACCATGATTCAGTGTGTGCCATACGACGG 415

QY 129 alProLeuLeuLeuAlaGlnTyrSerHisTyrPheGlnLysAlaArgAspThrLeuTyr 149
Db 416 TTCCCTCTTTTGTCAAAAACCTGATATTATTTTTCAGGAAGGCAATACACTATTATG 475

QY 149 eValTPrGlyGlnGlyArgHisMetAspArgValLeuGlyGlyArgThrTyrArgThr 169
Db 476 TTATTTGGGACCTTTCCGAAATATGAGAAAGATGGCAATGGCATGCTTACAACTGT 535

QY 169 eulLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyrThrPheThrGlyArgMet 189
Db 536 TGAATAAGACAGTGTGATCTATCCGAATGCCCAATATATGACACAGAGAGAGCGCA 595

QY 189 eValTyrCybAspGlnIlePheGlnAspGluThrGlyLysAsnArg----- 204

```

```

Db 596 TGAATTAAGTGTGATGAGATTTTAAGAGAAACTGGAGAGACAGTATTATCTACAC 655
QY 205 -----ArgGlnSerGlySerPheLeuSerThrGlyTyrP 216
Db 656 TGCATCAAAATCTTCCTCCAAATGATGTCAGTCTGGCTCATATCTCAGACAGGGGTGT 715
QY 216 heThMetIleLeuAlaLeuGlnLeuCybGlnGluIleValValTyrGlyMetValSer 236
Db 716 TTACTCTTCAATTCGGGCATGAGACGCTGTATGACATTCACGCTTACGGGATCAATTAATG 775
QY 236 spSerTyrCysArgGlyLysSerHisProSerValProTyrHisTyrPheGlnLysGly 256
Db 776 ACACCTTACTGACAGACAGAAAGGGTATGAAAGTCCCTTACATTTATTAACAAGAAA 835
QY 256 rglLeuAspGlyCysGlnMetTyrLeuAlaHisGlnGlnAlaProArgSerAlaHisArg 276
Db 836 GA---GATGAGTGTGATGATATATTTCTTCATGACATGCCCATATGGGGGTCAATG 892
QY 276 heIleThrGlyLysAlaValPheSerArgTPrAlaLysLysAspProIleValPheAla 296
Db 893 TTATCACTGAAAAGAAAGTGTTCCTAAATGGGCCAAAGACAGAGATATATTATCAC 952
QY 296 lAsProSerTyrP 299
Db 953 ATCCAAACTGG 963

RESULT 12
US-10-188-186-19
Sequence 19, Application US/10188186
Publication No. US20040029789A1
GENERAL INFORMATION:
APPLICANT: Anderson et al.
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-397C
CURRENT APPLICATION NUMBER: US/10/188,186
PRIOR FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: 60/303046
PRIOR FILING DATE: 2001-07-05
PRIOR APPLICATION NUMBER: 60/360814
PRIOR FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: 60/303828
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: 60/323380
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: 60/361133
PRIOR FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: 60/304016
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/304502
PRIOR FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: 60/305262
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: 60/373881
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/305673
PRIOR FILING DATE: 2001-07-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 368
SOFTWARE: Custom
SEQ ID NO 19
LENGTH: 1106
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (42)..(957)
US-10-188-186-19

Alignment Scores:
Pred. No.: 1,12e-76
Score: 684.00
Percent Similarity: 63.01%
Length: 1106
Matches: 132
Conservative: 52

```

Best Local Similarity: 45.21% Mismatches: 80
 Query Match: 42.07% Indels: 28
 DB: 16 Gaps: 6

US-10-019-114a-7 (1-302) x US-10-188-186-19 (1-1106)

```

QY 26 CysTrpAlaGlyLeuProLeuCysLeuAla----- 35
DB 106 TGGTGG-----TTGTGCGTCTGTAATGAAGAAATTCCTACCTACTA 153
QY 36 ThrCysLeuAsp-----HisHisPheProThrGlySerArgProThr 49
DB 154 ACTGCTTTGACAACTGCTGACAAAGTGATGATACCTCTCTACACATACA----- 204
QY 50 ValProGlyProLeuHisPheSer--GlyTyrSerSerValProAspGlyLysProLeu 69
DB 205 ---GGCGGCCCCCTTGCACATCCTCATGTGATGATTAATGTAAGAACACAGACCCCTTGC 261
QY 69 AlaArgLupProCysArgSerCysAlaValAlaSerSerGlyGlnMetLeuGlySer 89
DB 262 AACTGAC---TGACCTTTGTGTCATAGTCAAACTCAGGTCAAGATGTCGCGCAGA 318
QY 89 LysLeuGlyAlaGluHisPheSerAlaGluCysValPheArgMetAspGlnAlaProThr 109
DB 319 AGGTGGGAAATGAGATGATGATGATCTCTGCTGATTTGAGAAATGAAACATGCCCCACCA 378
QY 109 AlaGlyPheGluAlaAspValGlyGlnArgSerThrLeuArgValAlaSerHisTher 129
DB 379 AAGGTATGAAAGATGTCGCGCGCATGACATGATTCGAGTGTGTCATACCGCGC 438
QY 129 AlaProLeuLeuLeuArgSerHisTherHisTherPheGlnLysAlaArgAspThrLeu 149
DB 439 TTCCCTCTTTGCTTAAAAAACCCTGATTAATTTTCAAGGAAGCAATACTACTATTAATG 498
QY 149 eValLTPGlyGlnGlyArgHisMetAspArgValLeuGlyGlyArgThrTyrArgThr 169
DB 499 TTATTTGGGACCTTCCGCAATATGAGAAAGATGCAATGCGATGCTTACCAACATGT 558
QY 169 euleGlnLeuThrArgMetTyrProGlyLeuGlnValTyrThrPheThrGluArgMet 189
DB 559 TGAAGAAAGACAGTGTATCTATCCGAATGCCAAATATACGTACCAACAGAGAGCGCA 618
QY 189 eValAlaTyrCysAspGlnLeuPheGlnAspGluThrGlyLysAsnArgArgGlnSer 209
DB 619 TGAGTTACTGTGATGAGATGATTTTAAAGAAAGCTGGAGAACAGACATGACAGATGCG 678
QY 209 ePheLeuSerThrGlyTyrPheThrMetCileLeuAlaLeuGlnLeuCysGluGln 229
DB 679 TGTGATTGATCAAGAGGTGTTTACATTCATTCTGGCCATGAGCGCTGTATGGCATTC 738
QY 229 AlaValTyrGlyMetValSerAspSerTyrCysArgGluLysSerHisProSerValPro 249
DB 739 ACGTCTACGAGATGATTAATGACACCTACTGCAAGACAGAGGATATGAAAGTCCCT 798
QY 249 TyrHisTyrPheGlnLysGlyArgLeuAspGluCysGlnMetTyrLeuAlaHisGln 269
DB 799 ACCATTATATGAAACAGAGAGA---GATGAGTGTGATGATATTTCTTCTATGAACG 855
QY 269 LArProArgSerAlaHisArgPheIleThrGlyValAlaPheSerArgTrpAlaLys 289
DB 856 CCCCATATGAGGCGCATAGTTTATCACTGAAAGAAAGTGTGTTCTAAATGAGCGCAAGA 915
QY 289 LysArgProLleValPheAlaHisProSerTrp 299
DB 916 AGCAGAGATATATATTACACATCCAACTGG 947

```

RESULT 13

US-10-097-340-188
 ; Sequence 188, Application US/10097340
 ; Publication No. US20030087250A1
 ; GENERAL INFORMATION:
 ; APPLICANT: John MONAHAN
 ; APPLICANT: Manjula GANNANAVARAPU

```

/ APPLICANT: Sebastian HOERSCH
/ APPLICANT: Shudhangt KAMATKAR
/ APPLICANT: Steve G. KOVATS
/ APPLICANT: Rachel E. MEYERS
/ APPLICANT: Michael MORRISSEY
/ APPLICANT: Peter OLANDT
/ APPLICANT: Ami SEN
/ APPLICANT: Peter VEIRBY
/ APPLICANT: Gordon B. MILLIS
/ APPLICANT: Robert C. BAST, Jr.
/ APPLICANT: Karen LU
/ APPLICANT: Rosemarie SCHMANDT
/ APPLICANT: Xumei ZHAO
/ APPLICANT: Karen GIATT
/ TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
/ FILE REFERENCE: MRI-030
/ CURRENT APPLICATION NUMBER: US/10/097,340
/ PRIOR FILING DATE: 2002-03-14
/ PRIOR APPLICATION NUMBER: 60/276,025
/ PRIOR FILING DATE: 2001-03-14
/ PRIOR APPLICATION NUMBER: 60/325,149
/ PRIOR FILING DATE: 2001-09-26
/ PRIOR APPLICATION NUMBER: 60/276,026
/ PRIOR FILING DATE: 2001-03-14
/ PRIOR APPLICATION NUMBER: 60/324,967
/ PRIOR FILING DATE: 2001/09/26
/ PRIOR APPLICATION NUMBER: 60/311,732
/ PRIOR FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: 60/325,102
/ PRIOR FILING DATE: 2001-09-26
/ PRIOR APPLICATION NUMBER: 60/323,580
/ PRIOR FILING DATE: 2001-09-19
/ NUMBER OF SEQ ID NOS: 363
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 188
/ LENGTH: 2048
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-097-340-188

Alignment Scores:
Pred. No.: 1,32e-70 Length: 2048
Score: 639.50 Matches: 125
Percent Similarity: 64.75% Conservative: 44
Best Local Similarity: 47.89% Mismatches: 89
Query Match: 39.33% Indels: 3
DB: 14 Gaps: 3

US-10-019-114a-7 (1-302) x US-10-097-340-188 (1-2048)
QY 4 ThrdGlySerArgProThrValPro---GlyProLeuHisPheSerGlyTyrSerSerVal 62
DB 365 ACCGACAGCGGCCCGGGGATCCCGCGGACCGCGGCACTGAGCGATGCTCGGAGTG 424
QY 63 ProAspGlyLysProLeuValArgGluProCysArgSerCysAlaValAlaSerSer 82
DB 425 GCGGACCAAGCCCGG---AAAATGCACTGACGAGGAGCTGCTCGTGATGACACTCA 481
QY 83 GlyLysMetLeuGlySerGlyLeuGlyAlaGluHisPheSerAlaGluCysValPheArg 102
DB 482 GGGCATCTGTGCACTGCGCAAGGCTCCAGATGACCAAGACAGAGTGTGTCATCCGC 541
QY 103 MetAsnGlnAlaProThrValGlyPheGluAlaAspValGlyGlnArgSerThrLeuArg 122
DB 542 ATGAATACGCCGCCCAACGCGGTATGGCGGTGAGCTGGCAATGCGACACAGCTGAGG 601
QY 123 ValValSerHisThrSerValProLeuLeuLeuArgSerHisTherHisTherPheGln 142
DB 602 GTCATCCGCAATTCACAGATCCAGAGATCTCCGCAACCGCATGACCTGCTCAACGTG 661
QY 143 AlaArgAspThrLeuTyrMetValTTPGlyGlnGlyArgHisMetAspArgValLeuGly 162

```

Db 662 AGCAGGGCAACGCTTCTGAGGGCCGACGACATGCGGGGAGCGCAG 721
Qy 163 G1YArgThrTyrArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyr 182
Db 722 GGCACGCTCAACAACACTGCATCTCTGACGAGGTGCGCCGCGTGAAGGCTTC 781
Qy 183 ThrPheThrGluArgMetMetAlaTyrCyAspGlnIlePheGlnAspGluThrGlyLys 202
Db 782 ATGATTACTGCCCAAGATGCTGCACTTTGATGATGCTTCAAGCAGAGACTGCAAA 841
Qy 203 AsnArgGlnSerGlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeu 222
Db 842 GACAGGAAGATATCAACACTGCTGCTCAGACTGCTGTTTACAAAGACATTGCATG 901
Qy 223 GluLeuCyseGluGluIleValTyrGlyMetValSerAspSerTyrCyAspGluLys 242
Db 902 GAGCTGTGACAGATCAATGTTTATGCGATGCTGCCCCAGACTCTCGAGGATCCC 961
Qy 243 SerHisProSerValProTyrHisTyrPheGluLysGlyArgLeuAspGluCyseGlnMet 262
Db 962 AATCACCTTCAGACTTATCATTTATATGAACTTTTGACCTGATGAAATGTAATG 1021
Qy 263 TyrLeuAlaHisGluGlnAlaProArg---SerAlaHisArgPheIleThrGluValAla 281
Db 1022 TACTCTCCATGAGCAGAGCAGCAGCAGCATCATCACCCTTATCAGAGAAACGA 1081
Qy 282 ValPheSerArgTyrAlaLysLysArgProIleValPheAlaHisProSerTyrArgThr 301
Db 1082 GTCTTTAAGAACTGGCAGCAGCATTCATATTTCACTTTTTCACAGACTGAAACCA 1141
Qy 302 Glu 302
Db 1142 GAA 1144

RESULT 14

US-10-094-749-802
; Sequence 802, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: MAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 802
; LENGTH: 2090
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-802

Alignment Scores:

Pred. No.:	1,37e-70	Length:	2090
Score:	639.50	Matches:	125
Percent Similarity:	64.75%	Conservative:	44
Best Local Similarity:	47.89%	Mismatches:	89
Query Match:	39.33%	Indels:	3
DB:	15	Gaps:	3

US-10-019-114A-7 (1-302) x US-10-094-749-802 (1-2090)

Qy 44 ThrGlySerArgProThrValPro---GlyProLeuHisPheSerGlyTyrSerSerVal 62
Db 425 ACCCAGAGGCGCCCGGGGCTCCCGCGGACCCGCGCATGACGATACCTCGAGT 484
Qy 63 ProAspGlySerProLeuValArgGluProCyAspSerCyAlaValAlaSerSerSer 82
Db 485 GCGGACCAAGCCCTG---AAATGCACTGCAGGGAGCTGCGCTGTCGACAGCTCA 541
Qy 83 GlyGlnMetLeuGlySerGlyLeuGlyAlaGluIleAspSerAlaGluCyseValPheArg 102
Db 542 GGGCATCTGTGACAGCTCGCAAGGCTCCCAATGACAGACAGAGTGTCTATCCG 601
Qy 103 MetAsnGlnAlaProThrValGlyPheGluAlaAspValGlyGlnArgSerThrLeuArg 122
Db 602 ATGAATAGCCGCCACACCGCGCTATGCGGTGAGCGTGGCAATGACACAGCTGAG 661
Qy 123 ValValSerHisThrSerValProLeuLeuLeuArgHisTyrSerHisTyrPheGlnLys 142
Db 662 GTCATCGCATTCACAGCATCCAGAGATCTCCGCAACCGCATGACCTGCAACGCTG 721
Qy 143 AlaArgAspThrLeuTyrMetValTyrGlyGlnGlyArgHisMetLeuPheValLeuGly 162
Db 722 AGCCAGGGCACCGCTTCACTTCTGGGCGCCAGCAGCTACATCATCGCGCGGACG 781
Qy 163 GlyArgThrTyrArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyr 182
Db 782 GGCAGGCTCAACAACCTGCATCTCTGACGAGGTGCTGCCGCGTGAAGGCTTC 841
Qy 183 ThrPheThrGluArgMetMetAlaTyrCyAspGlnIlePheGlnAspGluThrGlyLys 202
Db 842 ATGATTACTGCCCAAGATGCTGCACTTTGATGAGCTCTTCAAGCAGAGACTGCAAA 901
Qy 203 AsnArgGlnSerGlySerPheLeuSerThrGlyTyrPheThrMetIleLeuAlaLeu 222
Db 902 GACAGGAAGATATCAACACTGCTGCTCAGCAGCTGCTGTTTCAATGACAATGCACTG 961
Qy 223 GluLeuCyseGluGluIleValTyrGlyMetValSerAspSerTyrCyAspGluLys 242
Db 962 GAGCTGTGACAGATCAATGTTTATGCGATGCTGCCCAAGACTTCTGAGGATCCC 1021
Qy 243 SerHisProSerValProTyrHisTyrPheGluLysGlyArgLeuAspGluCyseGlnMet 262
Db 1022 AATCACCTTCAGACTTATCATTTATATGAACTTTTGACCTGATGAAATGTAATG 1081
Qy 263 TyrLeuAlaHisGluGlnAlaProArg---SerAlaHisArgPheIleThrGluValAla 281
Db 1082 TACTCTCCATGAGCAGAGCAGCAGCAGCATCATCACCCTTATCAGAGAAACGA 1141
Qy 282 ValPheSerArgTyrAlaLysLysArgProIleValPheAlaHisProSerTyrArgThr 301
Db 1142 GTCTTTAAGAACTGGCAGCAGCATTCATATTTCACTTTTTCACAGACTGAAACCA 1201
Qy 302 Glu 302
Db 1202 GAA 1204

RESULT 15
US-10-288-252-31
; Sequence 31, Application US/10288252
; Publication No. US20030143686A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: LAL, Preeti G.

```

APPLICANT: TANG, Y. Tom
APPLICANT: YUE, Henry
APPLICANT: BURFORD, Neil
APPLICANT: GANDHI, Ameena R.
APPLICANT: WARREN, Bridget A.
APPLICANT: YAO, Monique G.
APPLICANT: TRIBOULET, Catherine M.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: LEE, Ernestine A.
APPLICANT: HAFALIA, April J.A.
APPLICANT: LU, Van
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: SANJANMALA, Madhu S.
APPLICANT: DING, Li
FILE OF INVENTION: TRANSFERASES
FILE REFERENCE: PI-0241 USA
CURRENT APPLICATION NUMBER: US/10/288, 252
CURRENT FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: PCT US01/30424
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US 60/252, 819
PRIOR FILING DATE: 2000-11-21
PRIOR APPLICATION NUMBER: US 60/249, 639
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: US 60/247, 931
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/246, 001
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US 60/244, 025
PRIOR FILING DATE: 2000-10-27
PRIOR APPLICATION NUMBER: US 60/238, 481
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/236, 523
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PERL Program
SEQ ID NO 31
LENGTH: 2075
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030143686A1 2778782CB1
US-10-288-252-31

```

```

Alignment Scores:
Pred. No.: 3,26e-70 Length: 2075
Score: 636.50 Matches: 124
Percent Similarity: 64.75% Conservative: 45
Best Local Similarity: 47.51% Mismatches: 89
Query Match: 39.15% Indels: 3
DB: 15 Gaps: 3

```

US-10-019-114A-7 (1-302) x US-10-288-252-31 (1-2075)

```

QY 44 ThrGlySerArgProThrValPro---GlyProLeuHisPheSerGlyTyrSerSerVal 62
DB 411 ACCGACGACGCGCCCGGGGTCCTCCGCGGACCGCGGCACATGACGATACCTCGAGATG 470
QY 63 ProAspGlyLysPProLeuValArgGluProCysArgSerCysAlaValValSerSerSer 82
DB 471 GCGGACCAAGAGCCCTG---AAATGCACTGCGAGGACTGTGCTGTGTGACCAAGCTCA 527
QY 83 GlyIleMetLeuGlySerGlyLeuGlyValGluIleLeuSerArgGluCysValPheArg 102
DB 528 GGGCATCTGTCGACAGTCCGCAAGGCTCCAGATTGACCAAGAGAGTGTGTCATCCGC 587
QY 103 MetLeuGlnAlaProThrValGlyPheGlnAlaAspValGlyGlnArgSerThrLeuArg 122
DB 588 ATGAATGACGCGCCCAAGCGGCTATGCGGTGACGTGGGCATTCGACCAAGCCTGAGG 647
QY 123 ValValSerHisThrSerValProLeuLeuLeuArgAsnTyrSerHisTyrPheGlnLys 142

```

```

DB 648 GTCATGCGCATTCAGCATCCAGAGATCTCCGCAACCGCCATGATGCTGCTCAACGTC 707
QY 143 AlaArgAspThrLeuTyrMetValTrrpglyGlnGlyArgHisMetAspArgValLeuGly 162
DB 708 AGCCAGGAGCCGCTGTCATCTTCTGGGCCCCAGACGCTACATGCGGGAGCGGCAAG 767
QY 163 GlyArgThrTyrArgThrLeuLeuGlnLeuThrArgMetTyrProGlyLeuGlnValTyr 182
DB 768 GCGCAGGTCTACACAACTGCATCTCCGAGCCAGGTCTGCGCGGCTGAAGGCTTC 827
QY 183 ThrPheThrGluArgMetLeuAlaTyrCysAspGlnIlePheGlnAspGlnThrGlyLys 202
DB 828 ATGATTACTCGCCACAAGATGCTGCGATTGATGAGCTTCAAGCAGAGAGACTGGCAAA 887
QY 203 AsnArgArgGlnSerGlySerPheLeuSerThrGlyTrrPheThrMetIleLeuAlaLeu 222
DB 888 GACAGAGAGATATCCAAACTTGGCTCAGACACTGGCTGTTTCAATGACATTCGACTG 947
QY 223 GluLeuCysGluGluIleValValTyrGlyMetValSerAspSerTyrCysArgGluLys 242
DB 948 GAGCTCTGACAGATCAATGTTTATGCGATGTCGCCCAAGACTTCTGCGAGGATCCC 1007
QY 243 SerHisProSerValProTyrHisTyrPheGluLysGlyArgLeuAspGluCysGlnMet 262
DB 1008 AATCACCCCTTCACTATCATTTATTAAGACCTTTTGAACCTGATGAATGTACAATG 1067
QY 263 TyrLeuAlaHisGluGlnAlaProArg---SerAlaHisArgPheIleThrGluLysAla 281
DB 1068 TACCTCTCCATGAGGAGGAGGACGCAAGGCGAGTCATCCGCTTTTTCACAGAAACGA 1127
QY 282 ValPheSerArgTrrPalaLysLysArgProIleValPheAlaHisProSerTrrParGthr 301
DB 1128 GTCTTTAAGACGTGGGACGACGACATTCATATTCATCTTTTTCACACGACTGGAACGA 1187
QY 302 Glu 302
DB 1188 GAA 1190

```

Search completed: December 22, 2004, 00:16:15
Job time : 610 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2004, 18:16:28 ; Search time 39 Seconds
(without alignments)
513.539 Million cell updates/sec

Title: US-10-019-114A-7
Perfect score: 1626
Sequence: 1 MKAPGRVLVLLICSVFSAV.....FSRWAKKRPVFAHPSWRTE 302

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A COMB .pep:*
2: /cgn2_6/ptodata/1/1aa/5B COMB .pep:*
3: /cgn2_6/ptodata/1/1aa/6A COMB .pep:*
4: /cgn2_6/ptodata/1/1aa/6B COMB .pep:*
5: /cgn2_6/ptodata/1/1aa/PTC05 COMB .pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	1626	100.0	302	4	US-09-599-360B-75
2	639.5	39.3	336	3	US-09-334-601-13
3	494	30.4	95	4	US-09-621-976-3915
4	198.5	12.2	340	1	US-08-446-875-16
5	196.5	12.1	359	1	US-08-503-133A-2
6	196.5	12.1	359	2	US-08-576-775A-2
7	196.5	12.1	359	2	US-08-972-498-2
8	196.5	12.1	359	2	US-08-899-545-2
9	191	11.7	343	2	US-08-446-875-2
10	191	11.7	343	2	US-08-102-385G-2
11	190.5	11.7	339	1	US-08-626-994A-3
12	190.5	11.7	339	3	US-08-957-742-3
13	190.5	11.7	364	1	US-08-626-994A-1
14	190.5	11.7	364	2	US-08-957-742-1
15	178	10.9	355	2	US-08-666-367B-6
16	178	10.9	355	2	US-08-143-438-6
17	178	10.9	356	2	US-08-666-367B-5
18	178	10.9	566	3	US-09-143-438-5
19	174.5	10.7	329	3	US-09-334-601-10
20	173.5	10.7	90	3	US-09-334-601-21
21	172	10.6	375	2	US-08-446-875-10
22	172	10.6	375	2	US-08-102-385G-10
23	171	10.5	376	2	US-08-666-367B-8
24	171	10.5	376	3	US-09-143-438-8
25	170	10.5	374	2	US-08-446-875-4
26	170	10.5	374	2	US-08-102-385G-4
27	168	10.3	90	3	US-09-334-601-20

28	168	10.3	406	1	US-08-446-777-4	Sequence 4, Appl
29	168	10.3	767	1	US-08-446-777-6	Sequence 6, Appl
30	168	10.3	767	1	US-08-446-777-8	Sequence 8, Appl
31	166.5	10.2	90	3	US-09-334-601-22	Sequence 22, Appl
32	166.5	10.2	90	3	US-09-334-601-24	Sequence 24, Appl
33	165	10.1	600	4	US-08-361-304A-2	Sequence 462, App
34	163.5	10.1	356	4	US-08-102-385G-18	Sequence 2, Appl
35	159	9.8	340	2	US-08-102-385G-18	Sequence 18, Appl
36	157.5	9.7	90	3	US-09-334-601-23	Sequence 23, Appl
37	148.5	9.1	90	3	US-09-334-601-25	Sequence 25, Appl
38	145.5	8.9	90	3	US-09-334-601-26	Sequence 26, Appl
39	143	8.8	329	1	US-07-991-587A-7	Sequence 7, Appl
40	143	8.8	329	1	US-08-309-985-7	Sequence 7, Appl
41	143	8.8	332	2	US-08-446-875-12	Sequence 12, Appl
42	143	8.8	332	2	US-08-102-385G-12	Sequence 12, Appl
43	143	8.8	384	3	US-09-334-601-8	Sequence 8, Appl
44	142.5	8.8	77	2	US-08-102-385G-29	Sequence 29, Appl
45	142	8.7	55	2	US-08-102-385G-15	Sequence 15, Appl

ALIGNMENTS

```
RESULT 1
US-09-599-360B-75
Sequence 75, Application US/09599360B
Patent No. 6548633
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Bougueleret, L.
TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
FILE REFERENCE: GENSET.050CP3
CURRENT APPLICATION NUMBER: US/09/599,360B
CURRENT FILING DATE: 2000-06-21
PRIOR APPLICATION NUMBER: 60/113,686
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/141,032
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 09/469,099
NUMBER OF SEQ ID NOS: 123
SOFTWARE: Patent.pm
SEQ ID NO 75
LENGTH: 302
TYPE: PRT
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -18..-1
US-09-599-360B-75
Query Match 100.0%; Score 1626; DB 4; Length 302;
Best Local Similarity 100.0%; Pred. No. 3.4e-170;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MKAPGRVLVLLICSVFSAVYIILCCAGCPPLATCTLDHFPPTGSRPTYPGLHSFGYS 60
1 LRVVSHTSVPLLRNTSHYFQKARDTLVYWGCGRHMDRVLAGRTYRTLLQLTRMPGLQ 180
121 LRVVSHTSVPLLRNTSHYFQKARDTLVYWGCGRHMDRVLAGRTYRTLLQLTRMPGLQ 180
121 LRVVSHTSVPLLRNTSHYFQKARDTLVYWGCGRHMDRVLAGRTYRTLLQLTRMPGLQ 180
181 VYFTEPMAYCCQIFQDETGNKRRSGSFLSGFTMLIALCEIIVVYGWVSDSYCR 240
181 VYFTEPMAYCCQIFQDETGNKRRSGSFLSGFTMLIALCEIIVVYGWVSDSYCR 240
241 EKSHPSVPHYFPEKGRLDEQMTLAHQAPRSARHPTTEKAVSRNAKKRPVFAHPSWR 300
```

Db 241 EXHPSVPIYIFKGRLEDECOMYLAHQAPRSARFITEKAVSRMAKRPPIFAHNSWR 300
QY 301 TE 302
Db 301 TE 302

RESULT 2

US-09-334-601-13
; Sequence 13, Application US/09334601
; Patent No. 6280989
; GENERAL INFORMATION:
; APPLICANT: Kapitonov, Dmitri
; APPLICANT: Yu, Robert
; TITLE OF INVENTION: NOVEL STALYITRANSFERASES
; FILE REFERENCE: VCUIP-6
; CURRENT APPLICATION NUMBER: US/09/334,601
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-334-601-13

Query Match 39.3%; Score 639.5; DB 3; Length 336;
Best Local Similarity 47.9%; Pred. No. 1,1e-61;
Matches 125; Conservative 44; Mismatches 89; Indels 3; Gaps 3;

QY 44 TGSRPYTP-GPLHSGYSSVPDCKEIVREPCRGAVVSSGQMLGSGLAGAIDSAECVFR 102
Db 64 TQGPVGPAGRPIDGLGVADHKPL-KMCRDCALVTSGLHSHQSQIDQTECVIR 122
QY 103 MNQAPYGPADYVQGRSTLVNVSHTSVPLILNYSHYFOKARDTLVWVGQGRHMDRVLG 162
Db 123 MNDAPYTGVRDVGNRSTLRVIAHNSIQRLIRNHDLLANSQGVFLFWGSSYMRDQK 182
QY 163 GRVYRLLQITRMVPGIQTFTERRMAVCDQIFQDGTGKRRSGSFLSTGFTMLAL 222
Db 183 GQVNNHLLSQVLPRLKAFWITRHKMLQFDELFOKETGDKRKISNTWLSTGFTMTIAL 242
QY 223 ELCEIIVYGVNSYGRKSHSPVPHYFEKGRLEDECOMYLAHQAPR-SARFITEKA 281
Db 243 ELCDRIIVYGVNPPDFCRDPHPSVPIYIIEPSPDECIMLSHERGRKSHRFFITEK 302
QY 282 VFSWAKKRPVFAHNSWRTE 302
Db 303 VFKWARTFNIHFQDPWKDE 323

RESULT 3

US-09-621-976-3915
; Sequence 3915, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.Dm
; SEQ ID NO 3915
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -18...-1

NAME/KEY: UNSURE
LOCATION: 59
OTHER INFORMATION: Xaa = Ala,Thr
NAME/KEY: UNSURE
LOCATION: 56
OTHER INFORMATION: Xaa = Arg,Ser
NAME/KEY: UNSURE
LOCATION: 54
OTHER INFORMATION: Xaa = His,Pro
US-09-621-976-3915

Query Match 30.4%; Score 494; DB 4; Length 95;
Best Local Similarity 96.8%; Pred. No. 1,7e-46;
Matches 92; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAPGLVLIILCSVFSNAVYIILCCWAGLPICLATCDHHPTGSRPTVPGPLHRSQYS 60
Db 1 MKAPGLVLIILCSVFSNAVYIILCCWAGLPICLATCDHHPTGSRPTVPGPLHRSQYS 60
QY 61 SVPDGKPLVREPCRGCAVWSSSGQMLGSGLAGAID 95
Db 61 SVPDGKPLVREKXCXCVWSSSGQMLGSGLAGAID 95

RESULT 4

US-08-446-875-16
; Sequence 16, Application US/08446875
; Patent No. 5838751
; GENERAL INFORMATION:
; APPLICANT: Paulson, James C.
; APPLICANT: Wen, Xiaohong
; APPLICANT: Livingston, Brian Duane
; APPLICANT: Gillespie, William
; APPLICANT: Kelm, George
; APPLICANT: Burlingame, Alma L.
; APPLICANT: Medzihradzky, Katalin
; TITLE OF INVENTION: Compositions and Methods for the
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poma, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; City: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,875
; FILING DATE: July 12, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/102,385
; FILING DATE: August 4, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 111-197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-446-875-16

Query Match 12.2%; Score 198.5; DB 2; Length 340;
Best Local Similarity 27.8%; Pred. No. 3,6e-13;
Matches 55; Conservative 32; Mismatches 60; Indels 51; Gaps 6;

QY 73 CRSCAVSSSGQMLGSLGAIEIDSAECVFRMNOAPTVGFADVGQRSTLRVSHTSVPLL 132
DB 139 CRCAVAVNGNGLRESYGPEDISHDFVLRNNKAPTAFTGFADVGTCTHLLV----- 190
QY 133 LRNYSHYFQKARPTLVVWVGQRHMD-----RVLGRTYRLL-----QUTRM 175
DB 191 ---YPSFRILGDNVSMILVFPKTLDEWVSAITGTISHITYIPVPAKIRVKODKILLY 247
QY 176 YPGIQTFTFRMAYCDQIFODETKNRQSSFLSTGFTMTLALCEIIVVGVMS 235
DB 248 HPAPIKTVFDMWLOGH-----GRYPTGLISVIFSNHVCDEVLLYFGA 291

QY 236 DSYCREKSHPSVPYHYFE 253
DB 292 DS--KGNMH-----HYME 302

RESULT 5
US-08-503-133A-2
; Sequence 2, Application US/08503133A
; Patent No. 5747326
; GENERAL INFORMATION:
; APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;
; TITLE OF INVENTION: Isolated Polysialyl Transferases,
; TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of
; TITLE OF INVENTION: Production and Use
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Felle & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/503,133A
; FILING DATE: 17-JULY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/04289
; FILING DATE: 22-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5747326man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: BOER 1050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-503-133A-2

Query Match 12.1%; Score 196.5; DB 1; Length 359;
Best Local Similarity 27.5%; Pred. No. 6,4e-13;
Matches 69; Conservative 39; Mismatches 98; Indels 45; Gaps 9;

QY 59 YSSVPGDKPLVREBRCSCAVVSSSGQMLGSLGAIEIDSAECVFRMNOAPTVGFADVGQR 118
DB 125 HSLLPVSPMKNRFRKTCVAVVNGSGLIDSGCGKEIDSHNFVLRCLNAPVVERFADVGTK 184

QY 119 STLRVSHTSVPLL-----LRVYS-----HYFQKARPTL-----YMWVGQRHMDRVLGGR 164
DB 185 SDFITMNPVSVOQAFGGFRNBSRDAKFVHRLSMLNDSVLMIPAPWVAGGRHVEW-NAL 243
QY 165 TYRLLQLTMYPGLOV-----YTFERMAYCDQIFODETKNRQSSFLSTGFTM 218
DB 244 ILNKLKRVTAIVYSLRIHVRGVWLTNNKVI-----KRP-----STGLLMY 285
QY 219 ILALICEIIVVGVMSDSYCREKSHPSVPYHYFEKGRUDECOMYLAHQAPSAHRFIT 278
DB 286 TLATRFCEDEIHLGFW--PPPKDLNGKAVKXHYD---DLKRYFSNAPSHPMPLEFKT 339

QY 279 EKAVFSRWAK 289
DB 340 LNVLHNKGALK 350

RESULT 6
US-08-576-775A-2
; Sequence 2, Application US/08576775A
; Patent No. 5849904
; GENERAL INFORMATION:
; APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;
; TITLE OF INVENTION: Isolated Polysialyl Transferases,
; TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of
; TITLE OF INVENTION: Production and Use
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Felle & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/576,775A
; FILING DATE: 21-December-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/503,133
; FILING DATE: 17-JULY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/04289
; FILING DATE: 22-December-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5849904man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: BOER 1050.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-576-775A-2

Query Match 12.1%; Score 196.5; DB 2; Length 359;
Best Local Similarity 27.5%; Pred. No. 6,4e-13;
Matches 69; Conservative 39; Mismatches 98; Indels 45; Gaps 9;

QY 59 YSSVPGDKPLVREBRCSCAVVSSSGQMLGSLGAIEIDSAECVFRMNOAPTVGFADVGQR 118
DB 125 HSLLPVSPMKNRFRKTCVAVVNGSGLIDSGCGKEIDSHNFVLRCLNAPVVERFADVGTK 184

QY 119 STRVSHTSVPL--LNYS-----HYOKARDTL-----YMWGGRHMDRYLGR 164
DB 185 SDFITNPSVYORAFGGFERNESDRAKVFHRLSLMINDSVLMI PAFWVGKGRHVEW-NAL 243
QY 165 TYRTLLQTLTMYPGLOV-----YPTERRMAVCDQIFODETKNRROSGSFLSTGWTM 218
DB 244 ILKNLKVATYPSLRILHNAVGYWLTNKVPI-----KRP-----STGLMW 285
QY 219 ILALICEBIVVGMVSDSYCREKSHPSVPHYFEKGRLDECOMYLAHQAPSAHREFT 278
DB 286 TLATRFCDLHLYGFW--PPKDLNGKAVKHYHD---DLKRYFSNAPRHPMLPEFT 339
QY 279 EKAVFSRMAK 289
DB 340 LNVLHNRGALK 350

RESULT 7
US-08-972-498-2
Sequence 2, Application US/08972498
Patent No. 5959078

GENERAL INFORMATION:
APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;
INVENTOR: Nakayama, Jun; Eckhardt, Matthias
TITLE OF INVENTION: Isolated Polysialyl Transferases,
TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of
TITLE OF INVENTION: Production and Use
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/972,498
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/576,775
FILING DATE: 21-December-1995
APPLICATION NUMBER: 08/503,133
FILING DATE: 17-July-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/04289
FILING DATE: 22-December-1994
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5959078man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: BOER 1050.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-972-498-2

Query Match 12.1%; Score 196.5; DB 2; Length 359;
Best Local Similarity 27.5%; Pred. No. 6,4e-13;
Matches 69; Conservative 39; Mismatches 98; Indels 45; Gaps 9;
59 YSSVDPGKPLVREPCRSQAVSSSGOMLGSGLGAIDSAECVFRMNOAPTVGFADVGQR 118

DB 125 HSLLPEVSPMKRRFRTCAVNGNSGILLDSGCCKEIDSHNFVIRCNLA PVEFAADVGR 184
QY 119 STRVSHTSVPL--LNYS-----HYOKARDTL-----YMWGGRHMDRYLGR 164
DB 185 SDFITNPSVYORAFGGFERNESDRAKVFHRLSLMINDSVLMI PAFWVGKGRHVEW-NAL 243
QY 165 TYRTLLQTLTMYPGLOV-----YPTERRMAVCDQIFODETKNRROSGSFLSTGWTM 218
DB 244 ILKNLKVATYPSLRILHNAVGYWLTNKVPI-----KRP-----STGLMW 285
QY 219 ILALICEBIVVGMVSDSYCREKSHPSVPHYFEKGRLDECOMYLAHQAPSAHREFT 278
DB 286 TLATRFCDLHLYGFW--PPKDLNGKAVKHYHD---DLKRYFSNAPRHPMLPEFT 339
QY 279 EKAVFSRMAK 289
DB 340 LNVLHNRGALK 350

RESULT 8
US-08-899-545-2
Sequence 2, Application US/08899545
Patent No. 6020201

GENERAL INFORMATION:
APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;
INVENTOR: Nakayama, Jun; Eckhardt, Matthias
TITLE OF INVENTION: Isolated Polysialyl Transferases,
TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor, Methods of
TITLE OF INVENTION: Production and Use
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,545
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/503,133
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6020201man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: BOER 1050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-899-545-2

Query Match 12.1%; Score 196.5; DB 3; Length 359;
Best Local Similarity 27.5%; Pred. No. 6,4e-13;
Matches 69; Conservative 39; Mismatches 98; Indels 45; Gaps 9;
59 YSSVDPGKPLVREPCRSQAVSSSGOMLGSGLGAIDSAECVFRMNOAPTVGFADVGQR 118
DB 125 HSLLPEVSPMKRRFRTCAVNGNSGILLDSGCCKEIDSHNFVIRCNLA PVEFAADVGR 184

QY 119 SLRVVSHTSVPL--LRNYS-----HYFOKARDTL-----YMWGGRHMDRVLGGR 164
DB 185 SDITMPSVVOAFGFGRHESBPAKFVRLHSLMNDVSLMIPAMVAGGKHWV--NAL 243
QY 165 TYRLLQLTRMYPGLOV-----YTFERMMAYCDOI FODETGKRRRQSGSFLSTGFTM 218
DB 244 ILNKKLKVRAVPSRLILHVRGYWLTNNKVP I-----KRP-----STGLLMY 285
QY 219 ILALSELCEIIVYGMVSDSCREKSHPSVPHYFEKRLDECOMYLAHEQAPRAHFFIT 278
DB 286 TLATRFCDIEIHLGYFW--PPPKDLNGKAVKHYHYD---DLKRYFSPNASPHRMPLFEKT 339
QY 279 EKAVFSRWAKK 289
DB 340 LNVLHNRGALK 350

RESULT 9
US-08-446-875-2
Sequence 2, Application US/08446875
Patent No. 5858751
GENERAL INFORMATION:
APPLICANT: Paulson, James C.
APPLICANT: Wen, Xiaohong
APPLICANT: Livingston, Brian Duane
APPLICANT: Gillespie, William
APPLICANT: Kelm, Sorge
APPLICANT: Burlingame, Alma L.
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Identification and Synthesis of Sialyltransferases
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Poma, Smith, Lande & Rose
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,875
FILING DATE: July 12, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/102,385
FILING DATE: August 4, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Oldenkamp, David J.
REGISTRATION NUMBER: 29,421
REFERENCE/DOCKET NUMBER: 111-197
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 343 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-446-875-2

Query Match 11.7%; Score 191; DB 2; Length 343;
Best Local Similarity 30.0%; Pred. No. 2.4e-12;
Matches 57; Conservative 36; Mismatches 77; Indels 20; Gaps 5;

QY 67 PLVRE---PCRSCAVSSSGQMGSLGAEIDSAECVFRNNOAPTVGFEADVQGRSTLRV 123
DB 133 PLEKRLVSCRCACVAVNGSNLKSYYGPDIDSHDFVLNNKAPTEGFEADVGSKTTHHF 192

QY 124 VSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGGRTYRLLQLTRMYPGLOVYT 183
DB 193 VYPSFRELQAEVS-----MILVPEKTTDLNVTLSATTGTTISHTYVPPA-KIKV 242
QY 184 FTERMMAYCDOI FODETGKRRRQSGSFLSTGFTMLALCEIIVYGMVSDSCREKS 243
DB 243 KKEKILYHPAFIKYVDFRMLQGHGRPSTGISTVIFSLSHICDEVDLYFGADS--KGNW 300
QY 244 HPSVPHYFE 253
DB 301 H-----HYWE 305

RESULT 10
US-08-102-385G-2
Sequence 2, Application US/08102385G
Patent No. 5962294
GENERAL INFORMATION:
APPLICANT: Paulson, James C.
APPLICANT: Wen, Xiaohong
APPLICANT: Livingston, Brian Duane
APPLICANT: Gillespie, William
APPLICANT: Kelm, Sorge
APPLICANT: Burlingame, Alma L.
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Identification and Synthesis of Sialyltransferases
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oppenheimer Wolff & Donnelly LLP
STREET: 2029 Century Park East, 38th Floor
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/102,385G
FILING DATE: 04-AUG-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/925369
FILING DATE: 04-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Oldenkamp, David J.
REGISTRATION NUMBER: 29,421
REFERENCE/DOCKET NUMBER: 97-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 343 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-102-385G-2

Query Match 11.7%; Score 191; DB 2; Length 343;
Best Local Similarity 30.0%; Pred. No. 2.4e-12;
Matches 57; Conservative 36; Mismatches 77; Indels 20; Gaps 5;

QY 67 PLVRE---PCRSCAVSSSGQMGSLGAEIDSAECVFRNNOAPTVGFEADVQGRSTLRV 123
DB 133 PLEKRLVSCRCACVAVNGSNLKSYYGPDIDSHDFVLNNKAPTEGFEADVGSKTTHHF 192
QY 124 VSHTSVPLLRNYSHYFOKARDTLVWVGGRHMDRVLGGRTYRLLQLTRMYPGLOVYT 183

Db 193 VPESEFELAQEVS-----MILVFKTDLIEWISATTTGTSHTYVPA-KIKV 242

QY 184 FTERMAVCOIQIDGDKRKROSGSFLSTGFTMTLALCEIIVYGWSDSYCREKS 243

Db 243 KKEKILHPAFIKYVDRLWLOHGRIPTGISTSVISFLHICBVDLYRGAGS--KGMW 300

QY 244 HPSVPHYFE 253

Db 301 H-----HYWE 305

RESULT 11

US-08-626-994A-3

Sequence 3, Application US/08626994A

Patent No. 5798244

GENERAL INFORMATION:

APPLICANT: Shuchi TSUJI et al.

TITLE OF INVENTION: Sta' 2,3gaia 1,4GICNAC ' 2,8-

TITLE OF INVENTION: STALYTRANSPERASE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Greenblum & Bernstein, P.L.C.

STREET: 1941 Roland Clarke Place

CITY: Reston

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 20191

COMPUTER READABLE FORM:

MEDIUM TYPE: 1.44 diskette

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Word Perfect 5.1+ (ASCII)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/626,994A

FILING DATE: April 3, 1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 77469/1995

FILING DATE: April 3, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Arnold Turk

REGISTRATION NUMBER: 33,094

REFERENCE/DOCKET NUMBER: P14595

TELEPHONE: (703) 716-1191

TELEFAX: (703) 716-1180

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 339 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

US-08-626-994A-3

Query Match 11.7%; Score 190.5; DB 1; Length 339;

Best Local Similarity 30.0%; Pred. No. 2,7e-12;

Matches 65; Conservative 28; Mismatches 81; Indels 43; Gaps 8;

QY 60 SSVPDGKPLVREPCSCAVAVSSSGMIGSLGAEIDSAECVFRMNOAFTVGFADVGORS 119

Db 105 SLDPVSPIMKRYNVCAYVNSGILTGSCGQEIIXSDVSRCNPAFTEAFHKDVGRT 164

QY 120 TLRVVSHTSVPLLRNYSHYFOKARDTL-----YMWGQGRHMDRVLGRTYRT 168

Db 165 NLT-----TFNPISILEKYNMLLTIODRNPFLLSKLDGAILM-----IPAFPH 211

QY 169 LLOLTRYMYPGLQVYTFER-----MMAYCQIFQ--DETGKRRROSGSFLSTGFTMTL 220

Db 212 SATVTRTL-----VDFVEHKGQLKVQLAMPNIMQHVNRKYNKSKLSPKRLSTGILMYTL 267

QY 221 ALBLCEIIVYGWVS--DSYCREKSHPSVPHYFEK 254

Db 268 ASAICEIHLHYGFWPFGDPNTR-----DLPYHYDK 300

RESULT 12

US-08-957-742-3

Sequence 3, Application US/08957742

Patent No. 6017743

GENERAL INFORMATION:

APPLICANT: Shuchi TSUJI et al.

TITLE OF INVENTION: Sta' 2,3gaia 1,4GICNAC ' 2,8-

TITLE OF INVENTION: STALYTRANSPERASE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Greenblum & Bernstein, P.L.C.

STREET: 1941 Roland Clarke Place

CITY: Reston

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 20191

COMPUTER READABLE FORM:

MEDIUM TYPE: 1.44 diskette

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Word Perfect 5.1+ (ASCII)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/957,742

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/626,994

FILING DATE: April 3, 1996

APPLICATION NUMBER: JP 77469/1995

FILING DATE: April 3, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Arnold Turk

REGISTRATION NUMBER: 33,094

REFERENCE/DOCKET NUMBER: P14595

TELEPHONE: (703) 716-1191

TELEFAX: (703) 716-1180

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 339 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

US-08-957-742-3

Query Match 11.7%; Score 190.5; DB 3; Length 339;

Best Local Similarity 30.0%; Pred. No. 2,7e-12;

Matches 65; Conservative 28; Mismatches 81; Indels 43; Gaps 8;

QY 60 SSVPDGKPLVREPCSCAVAVSSSGMIGSLGAEIDSAECVFRMNOAFTVGFADVGORS 119

Db 105 SLDPVSPIMKRYNVCAYVNSGILTGSCGQEIIXSDVSRCNPAFTEAFHKDVGRT 164

QY 120 TLRVVSHTSVPLLRNYSHYFOKARDTL-----YMWGQGRHMDRVLGRTYRT 168

Db 165 NLT-----TFNPISILEKYNMLLTIODRNPFLLSKLDGAILM-----IPAFPH 211

QY 169 LLOLTRYMYPGLQVYTFER-----MMAYCQIFQ--DETGKRRROSGSFLSTGFTMTL 220

Db 212 SATVTRTL-----VDFVEHKGQLKVQLAMPNIMQHVNRKYNKSKLSPKRLSTGILMYTL 267

QY 221 ALBLCEIIVYGWVS--DSYCREKSHPSVPHYFEK 254

Db 268 ASAICEIHLHYGFWPFGDPNTR-----DLPYHYDK 300

RESULT 13

US-08-626-994A-1

Sequence 1, Application US/08626994A

Patent No. 5798244

GENERAL INFORMATION:

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/666,367B
FILING DATE: August 19, 1996

CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER:
FILING DATE:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ernest M. Choate, Jr.
FIRM: Choate, O'Connell & Smith, P.C.
ADDRESS: 1000 Massachusetts Avenue, N.W.
CITY: Washington, D.C. 20004
STATE: D.C.
COUNTRY: U.S.A.
TELEPHONE: (202) 638-1000
FAX: (202) 638-1000
E-MAIL: ernest.choate@choate.com
FIRM E-MAIL: info@choate.com

NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367

REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-8850
TELEFAX:

TELEFAX:
TELEX:
INFORMATION FOR CEN ID NO. 6:

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:

LENGTH: 355 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear

TOPOLOGI: linear
8-666-367B-6

every Match 10.9%;

st Local Similarity 25.2%; F
tches 63; Conservative 37;

73 CRSCAVSSSGQMLGSLGAE

```

      /5 CKSCLAVSSSGQMLDGGGLGAL
          ||||| : | : ||:| |
126 CTSCAVVGNGGITLNNSGMGQOE

```

126 CISC AVGNGLNNSGMGE
133 FNRGIV EGRADDEI YATWQ

133 LRNYSHY-EQKARDTLYMWG
|:| | |:

186 LQNLGHKGFKKI-----E

182 -----YTFTE
:|:

231 FLNYYGRPRERFDEFTMNH

Query Match	Score	DB 2;	Length
10.9%	178;	DB 2;	355;

Best Local Similarity 25.2%; Pred. No. 6.8e-11;
Matches 63; Conservative 37; Mismatches 80; Indels 70; Gaps 9.

73 CRSCAVSSSGOMLGSLGAHIDSAECVFRNQAPTVGFEADVGRSTLRVVSHTSVPL 132

126 TCTCAATGCGGATTTNNSGMSGQETDSDHYEVRVSGAVIKYGEKQDVGTSTSFYGETAASLYSS 185

```

126 CISCVAWNGGLINNSGWEIDSHDVFVRSVGAVIINGIENLVGINLSTFGRFALISLVWS 103
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111

```

```

133 LNRYSH-FQKARDLYMWGQGGRHMDVLGGRTYRLLQLTRMYSGLQV-----181
      |   |   |   |   |   |   |   |   |   |   |   |   |   |

```

186 LÖNLGHKGFKI-----PQKHI-----RYHFLAVRDYEWLKALLDDKDIRG 230

```

182 -----YFTTERMMAYCDQIFQDETGNRRQSGL-----STGW 215
          :|:::|:|||||:|

```

231 FLNYYGRRPRERFDEDFTNKKYLVAHDFL--RYLKNRFLSKNULOQPYWRLYRPTTGA 287

216 FTMLALELCEETIVVYGVWSDSYCREKSHPSVPYHYFEKGRLDECOMYLAHEQAQPSAHR 275

```

288 LLLLTALHLCDRYSAYGYITE-----GHQKSDHYIDK-EMKRLVEYVNH----- 332

```

276 FITEKAVESSR 285

2/6 FILENAVFSK 285
233 ENI PROTOYD 242

Search completed: December 10, 2004, 18:26:42
Job time : 40 secs

Job time : 40 secs